

# RELEASE

(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mparcn\_n n.e. n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 21:30:34 1999; Maspar time 353.20 Seconds

Tabular output not generated. 921.792 Million cell updates/sec

Title: >US-08-946-869-9  
Description: (1-2393) from US08946869.seq

Perfect Score: 2393 1 AAGGCCACCTCCGCCGCTCTC.....AAAAAAGCGCGCGCGGT 2393

N.A. Sequence: Comp: TTCGGGTGGAGCGCGAGAG.....TTTTTCCCGCGCGCGCA

Scoring table: TABLE default

Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0

Listing first 45 summaries

Database:

n:geneseq2  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics: Mean 9.947; Variance 7.935; scale 1.253

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	116	4.8	116	22	T25991	Human gene signature	1.45e-41
2	48	2.0	91	9	Q51746	Oligonucleotide probe	6.92e-09
3	45	1.9	204	1	N81164	Base substituted E.co	1.44e-07
4	43	1.8	91	9	Q51746	Oligonucleotide probe	1.05e-06
5	42	1.8	204	1	N81164	Base substituted E.co	2.85e-06
6	40	1.7	114	12	Q70468	Generic DNA sequence	2.03e-05
7	40	1.7	172	32	T76363	Human interleukin 8 a	2.03e-05
8	40	1.7	946	13	Q80219	Human NDF-alpha2b clo	2.03e-05
9	41	1.7	1098	13	Q80225	Rat NDF clone 4 DNA.	2.03e-05
10	40	1.7	1651	13	Q80218	Human NDF-alpha2b clo	2.03e-05
11	40	1.7	1734	33	T39792	Human SHP9 gene.	2.03e-05
12	40	1.7	3344	13	Q80228	Rat NDF clone 22 DNA.	2.03e-05
13	40	1.7	3451	38	V02308	Cell membrane proton-	2.03e-05

14	38	1.6	114	12	Q70466	Generic DNA sequence	1.40e-04
15	39	1.6	178	32	T76405	Human endothelin-1 an	3.35e-05
16	38	1.6	369	38	V02169	Human secreted protei	1.40e-04
17	38	1.6	369	39	T88086	3' portion of cDNA cl	1.40e-04
18	38	1.6	1425	23	T35230	Cytoplasmic antiprote	1.40e-04
19	39	1.6	2335	13	Q80216	Human prointra-alpha2b	3.35e-05
20	38	1.6	2400	1	Q05055	Placenta-specific pr	1.40e-04
21	38	1.6	2862	14	Q84612	Mouse A12 receptor cl	1.40e-04
22	36	1.5	56	9	Q52732	Sequence of oligo nuc	9.40e-04
23	36	1.5	114	12	Q70467	Generic DNA sequence	9.40e-04
24	36	1.5	114	12	Q70469	Generic DNA sequence	9.40e-04
25	36	1.5	114	12	Q70470	Generic DNA sequence	9.40e-04
26	36	1.5	114	12	Q70467	Generic DNA sequence	9.40e-04
27	36	1.5	162	32	T76307	Human RANTES antisens	9.40e-04
28	36	1.5	201	38	V00418	3' fragment of clone	9.40e-04
29	36	1.5	607	38	V02138	Human secreted protei	9.40e-04
30	36	1.5	607	39	T97397	Human secreted protei	9.40e-04
31	36	1.5	607	39	T88059	Partial cDNA clone en	9.40e-04
32	36	1.5	701	32	T77001	Mouse PLA2s cDNA (B6a	9.40e-04
33	36	1.5	810	32	T48835	Mouse PLA2s cDNA (w11	9.40e-04
34	37	1.5	857	22	T14564	Human milk kappa-case	3.64e-04
35	37	1.5	857	8	Q46850	Recombinant human kap	3.64e-04
36	37	1.5	906	13	Q80222	Human NDF-beta3 clone	3.64e-04
37	37	1.5	1338	11	Q65607	Rabbit zona pellucida	3.64e-04
38	37	1.5	1348	12	Q78278	T. nlyeum GAPDH.	9.40e-04
39	36	1.5	1502	39	V10099	Human MAP kinase p38-	9.40e-04
40	36	1.5	1737	17	Q99007	Clemokine superfamily	9.40e-04
41	37	1.5	2611	31	T67231	DNAK accessory molecu	3.64e-04
42	36	1.5	2884	11	Q84613	Rat A12 receptor clon	9.40e-04
43	37	1.5	3581	10	Q44391	Sequence of murine OS	3.64e-04
44	36	1.5	4091	1	Q04035	Stem cell leukaemia (	9.40e-04
45	36	1.5	4199	1	Q05330	Stem cell leukaemia (	9.40e-04

## ALIGNMENTS

RESULT 1  
ID T25991 standard; cDNA to mRNA; 116 BP.  
AC T25991;  
DE 28-OCT-1996 (first entry)  
DE Human gene signature H00508226.  
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KW human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; 88.  
OS Homo sapiens.  
PN M09514772-A1.  
PD 01-JUN-1995.  
PE 11-NOV-1994; J01916.  
PR 12-NOV-1993; JP-355504.  
PA (MATSU) MATSUBARA K.  
PA (OXUBO) OKUBO K.  
PI Matsubara K, Okubo K;  
DR MPI; 95-206931/27.  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissues  
PS Claim 1; Page 19/6; 2245bp; Japanese.  
CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC given in T19001-T26837 and which is able to hybridise to part of  
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(7) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (seep. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.

[illegible]

AC N81164;  
 DE 08-NOV-1990 (first entry)  
 DE Base substituted E.coli beta-galactosidase alpha-fragment.  
 KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.  
 OS Escherichia coli.  
 FH Key  
 FT misc\_feature  
 FT 19..69  
 FT /tag- a  
 FT /function-multiple cloning site  
 FT primer\_bind  
 FT 187..204  
 FT /\*tag- b  
 FT  
 FT EP-285123-A.  
 PD 05-MAY-1988.  
 PF 30-MAR-1988: 105163.  
 PR 03-APR-1987: US-034819.  
 PA (SUSO) SUOMEN SOKERT OY.  
 PI Lehtovara P, Knowles J, Kolvola A, Bamford J, Reinikainen T;  
 DR MPI: 88-279927/40.  
 PT Introducing random point mutations into nucleic acids -  
 PT by prepn of single stranded template, annealing a primer, elongation,  
 PT misincorporation, completion of molecules and screening.  
 PS Disclosure: P: English.  
 CC Random point mutations were introduced into the alpha fragment of  
 CC E.coli beta-galactosidase. The wild type sequence was obtained as a  
 CC single stranded template and an oligonucleotide was hybridised to  
 CC it to generate a popn of DNA molecules which terminate at all  
 CC possible nucleotide positions within a specified region. The  
 CC variable 3' ends generated in this way are used as primers for  
 CC reverse transcriptase. Nucleotides are misincorporated by the  
 CC transcribed and the molecules are completed to forms that can be  
 CC amplified and then expressed in a suitable host-vector system.  
 CC The sequence covers all 176 diff base substitutions, most of which  
 CC occurred singularly in any given mutant.  
 CC See also P80575.  
 CC Sequence 204 BP: 21 A; 47 C; 17 G; 11 T; 108 Others;  
 SO

Query Match 1.8%; Score 42; DB 1; Length 204;  
 Best Local Similarity 9.4%; Pred. No. 2.85e-06;  
 Matches 10; Conservative 57; Mismatches 39; Indels 0; Gaps 0;

Db 80 hvgcgymttcthyrmbyrdynrsdaawcyccyrsvrrsdcynachddhyvb 139  
 Oy 1324 TGGCCCTATGAGCAGCACCACCTGCTCGGAGCCGCTTTCGGACACCTCTG 1383

Db 140 bbyvvhnhnnccebnhvhnhvbnhnmwyrhdaiddv 185  
 Oy 1384 GGGCCAGCGGCGACACTGCTCTCCAGTATGCCAAGTGGCTGG 1429

RESULT 6  
 ID 070468 standard; DNA: 114 BP.  
 AC 070468;  
 DE 05-APR-1995 (first entry)  
 DE Generic DNA sequence to generate a random TSAR peptide library.  
 KW TSAR: totally synthetic affinity reagent; synthetic; binding domain;  
 KW effector domain; concatenated heterofunctional protein; linker;  
 KW direct; rapid; detection; screening; treatment; generic; ss.  
 OS Synthetic.  
 FH Key  
 FT misc\_feature  
 FT 55..60  
 FT /tag- a  
 FT /note- this sequence represents 'Z'; Z can be a  
 FT sequence of 6, 9 or 12 nucleotides (see  
 FT comments).  
 FT  
 FT MO9418318-A.  
 PD 18-AUG-1994.  
 PF 01-FEB-1994: US-009077.  
 PR 01-FEB-1993: US-013416.  
 PR 30-DEC-1993: US-176500.  
 PR 31-JAN-1994: US-189331.  
 PA (UYNC-) UNIV NORTH CAROLINA.  
 PI Fowler DM, Kay BK;  
 DR MPI: 94-279739/34.

DR P-PSDB: R65154.  
 PT Identifying proteins or peptide(s) which bind a ligand - by  
 PT screening a recombinant vector library expressing fusion proteins  
 PT comprising a binding domain and an effector domain  
 PS Disclosure: Page 35; 255pp; English.  
 CC 070468 is a generic DNA sequence used to generate random TSAR (Totally  
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
 CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNB)10Y. X  
 CC and Y are flanking restriction sites (X is not the same as Y) that are  
 CC not specified further. Other generic sequences are shown in 070466-68.  
 CC Other specific peptides generated by these generic sequences are shown in  
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
 CC comprising at least two functional regions - a binding domain with  
 CC affinity for a ligand and a second effector peptide portion that is  
 CC chemically or biologically active. They may further comprise a linker  
 CC peptide between the 2 domains. The oligonucleotides are also designed so  
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
 CC in, or flanking, the unpredicted or variant residues. These residues  
 CC confer some degree of conformational rigidity to the peptides. The TSARs  
 CC or compens, comprising a TSAR binding domain can be used in vivo to  
 CC deliver a chemically or biologically active moiety, eg. metal ion,  
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
 CC cell. They can also replace the function of macromolecules, eg.  
 CC monoclonal or polyclonal antibodies and therefore circumvent the need  
 CC for complex methods of hybridoma formation or in vivo antibody  
 CC production. The TSARs are easily characterised and have designed activity  
 CC allowing direct and rapid detection in a screening process.  
 CC Sequence 114 BP: 0 A; 2 C; 2 G; 2 T;  
 SO

Query Match 1.7%; Score 40; DB 12; Length 114;  
 Best Local Similarity 3.4%; Pred. No. 2.03e-05;  
 Matches 6; Conservative 34; Mismatches 72; Indels 0; Gaps 0;

Db 3 bnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnnnn 62  
 Oy 225 TGGGGGATGGGCGTGTACCTCTTGGCGAGCTCTATCGCAGAGTGGAGCGCTGC 284

Db 63 bnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnb 114  
 Oy 285 CCGCTCTCTCTCCGTCGTCGTCGCGCTCTGGGCGATGCCATTGCTCTCC 336

RESULT 7  
 ID T76363 standard; DNA: 172 BP.  
 AC T76363;  
 DE 15-SEP-1997 (first entry)  
 DE Human interleukin 8 antisense oligonucleotide.  
 KW Asthma; airway epithelium; adenosine free; cystic fibrosis;  
 KW chronic obstructive pulmonary disease; bronchitis; ss.  
 OS Synthetic.  
 FH MO9640162-A1.  
 PD 19-DEC-1996.  
 PR 06-JUN-1996; U09306.  
 PR 07-JUN-1995; US-474497.  
 PA (UYEC-) UNIV EAST CAROLINA.  
 PI Metzger WJ, Nyce JW;  
 DR MPI: 97-051871/05.  
 PT Treatment of airway diseases such as asthma - by topically applying  
 PT adenosine-free antisense oligo:nucleotide to airway epithelium of  
 PT subject  
 PS Claim 5; Page 36; 71pp; English.  
 CC A method for treating airway disease in a subject has been produced,  
 CC which involves the topical administration of an essentially adenosine  
 CC free antisense oligonucleotide (ON) to the airway epithelium of the  
 CC subject. The present sequence is an antisense oligonucleotide specific  
 CC for the human interleukin 8, targeted at the initiation codon. The  
 CC method can be used to treat airway diseases such as cystic fibrosis,  
 CC asthma, chronic obstructive pulmonary disease, bronchitis and other  
 CC airway diseases characterised by an inflammatory response. By  
 CC eliminating adenosine from the antisense ON, its liberation upon  
 CC antisense degradation is prevented, thereby preventing adenosine-  
 CC induced bronchoconstriction in patients with hyper-reactive airways.  
 CC Sequence 172 BP: 0 A; 35 C; 42 G; 39 T;

```

KW Alpha, beta; neu differentiation factor; NDF; human; rat; p185-neu;
KM tyrosine phosphorylation; differentiation; phenotype; proliferation;
KM wound; tumour; epithelial tissue; breast; stomach;
KM gastrointestinal disease; Barrett's oesophagus;
KM (non-)cystic kidney disease; inflammatory bowel disease; ds.
OS Rattus rattus.
FH Key
FT cdb
FT Location/Qualifiers
FT 17..712
FT /*tag= a
FT /product= Rat NDF
PN MO9428133-A.
PD 08-DEC-1994.
PE 23-MAY-1994. U05769.
PR 21-MAY-1993; US-066384.
PA (AMGE-) AMGEN INC.
PI Hu S, Koshl RA, Liu N, Pierce GF, Sugarman BJ;
PI Wen D,
PI WPI: 95-022805/03.
DR P-PSDB; R68570.
PT New recombinant neu differentiation factors and corresp. DNA -
PT are used in the treatment of tumours, dermal wounds, and
PT gastrointestinal, kidney and inflammatory bowel diseases.
PS Disclosure: Page 170-173, 341pp, English.
CC The sequences given in Q80223-93 encode-rat neu differentiation
CC factors (NDF 8). The peptides encoded by these cDNA clones isolated
CC from rat tissues and cell lines, possess the ability to stimulate
CC human p185-neu tyrosine phosphorylation. These peptides have the
CC ability to induce a differentiated phenotype in certain cell lines and
CC can stimulate or inhibit proliferation of certain cell lines. The
CC NDF's can be used to treat wounds, tumours derived from epithelial
CC tissue of the breast, stomach etc., gastrointestinal disease, Barrett's
CC oesophagus, (non-)cystic kidney disease or inflammatory bowel disease.
CC These cDNA sequences were used in the isolation of related sequences
CC from human cDNA libraries which encode human NDF's.
SQ Sequence 1098 bp; 343 A; 265 C; 254 G; 236 T;

Query Match
Best Local Similarity 1.7%; Score 41; DB 13; Length 1098;
Matches 53; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

DB 1034 tggtaataaataatcattccacaaagtataaaaaaaaaaaaaaaaaaagggcg 1093
QY 2326 TGGGGAAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAGGCGC 2385
DB 1094 cccgc 1098
QY 2386 GCCGC 2390

RESULT 10
ID 080218 standard; DNA; 1651 bp.
AC 080218:
DT 31-AUG-1995 (first entry)
DE Human NDF-alpha2b clone 17 DNA.
KW Alpha; beta; neu differentiation factor; NDF; human; rat; p185-neu;
KM tyrosine phosphorylation; differentiation; phenotype; proliferation;
KM wound; tumour; epithelial tissue; breast; stomach;
KM gastrointestinal disease; Barrett's oesophagus;
KM (non-)cystic kidney disease; inflammatory bowel disease; ds.
OS Homo sapiens.
FH Key
FT cdb
FT Location/Qualifiers
FT 20..830
FT /*tag= a
FT /product= Human NDF-alpha2b
PN MO9428133-A.
PD 08-DEC-1994.
PE 23-MAY-1994; U05769.
PR 21-MAY-1993; US-066384.
PA (AMGE-) AMGEN INC.
PI Hu S, Koshl RA, Liu N, Pierce GF, Sugarman BJ;
PI Wen D,
PI WPI: 95-022805/03.
DR P-PSDB; R68563.

```

PT New recombinant neu differentiation factors and corresp. DNA -  
 PT are used in the treatment of tumours, dermal wounds, and  
 CC gastrointestinal kidney and inflammatory bowel diseases.  
 PS Claim 25: Page 139-141: 341pp; English.  
 CC The sequences given in Q80215-22 encode human neu differentiation  
 CC factors (NDF's). The peptides encoded by these cDNA clones isolated  
 CC from human tissues and cell lines, possess the ability to stimulate  
 CC human p185-neu tyrosine phosphorylation. These peptides have the  
 CC ability to induce a differentiated phenotype in certain cell lines and  
 CC can stimulate or inhibit proliferation of certain cell lines. The  
 CC NDF's can be used to treat wounds, tumours derived from epithelial  
 CC tissues of the breast, stomach etc., gastrointestinal disease, Barrett's  
 CC oesophagus, (non-)cystic kidney disease or inflammatory bowel disease.  
 CC These cDNA sequences may be used in a recombinant plasmid for the  
 CC recombinant production of the NDF's in a pro- or eukaryotic host cell.  
 SO Sequence 1651 BP; 510 A; 460 C; 352 G; 349 T;

Query Match 1.7%; Score 40; DB 13; Length 1651;  
 Best Local Similarity 100.0%; Pred. No. 2.03e-05;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1612 aaaaaaaaaaaaaaaaaaaaaaaaaagggcgccgc 1651  
 ||||||||||||||||||||||||||||||||||||  
 OY 2351 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGC 2390

RESULT 11  
 ID T39792 standard; DNA; 1734 BP.  
 AC T39792;  
 DT 19-FEB-1998 (first entry)  
 DE Human SH3P9 gene.  
 KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;  
 KW cellular signalling element; cellular structural element; malignancy;  
 KW protein identification; functional domain; protein screening;  
 KW cellular signal transduction process; ss.  
 OS Homo sapiens.  
 FH Key  
 FT CDS Location/Qualifiers  
 FT 1..1215  
 FT /tag= a  
 FT /product= human SH3P9  
 PN MO9631625-A1.  
 PD 10-OCT-1996.  
 PE 04-APR-1996; U04454.  
 PR 03-APR-1996; US-630915.  
 PR 07-APR-1995; US-417872.  
 PA (CYTO-) CYTOGEN CORP.  
 PA (UYNC-) UNIV NORTH CAROLINA.  
 PI Fowlkes DM, Hoffman N, Kay BK, McConnell SJ, Sparks AB;  
 DR WPI: 96-465045/46.  
 DR P-PSDB: W05392.  
 PT Identifying polypeptide(s) having specific functional domain (esp.  
 PT SH3 domain) - comprises detecting selective binding to recognition  
 PT unit, regardless of sequence homology  
 PS Claim 76: Fig 36; 17pp; English.  
 CC T39786-T39803 represent novel human and mouse genes encoding Src-homology  
 CC region 3 (SH3) domain containing proteins that can be used in the method  
 CC of the invention. SH3 domain containing proteins play a role in  
 CC signalling and structural elements of cells. The method of the invention  
 CC is for identifying polypeptides containing functional domains of interest  
 CC (especially SH3 domains). The method comprises contacting a multivalent  
 CC recognition unit (RU) complex with a number of peptides and identifying  
 CC polypeptides having a selective binding affinity for the RU complex. The  
 CC method is based on functional similarities and does not rely on sequence  
 CC similarities. Prior methods only gave limited success for identifying  
 CC proteins which contain an SH3 domain due to the minimal sequence homology  
 CC among known SH3 proteins. It has been found that small peptide RUs in  
 CC multivalent form have reduced specificity for a given peptide RUs in  
 CC compared to monomer RUs. Multivalent RU complexes are particularly suited  
 CC to screening for polypeptides containing functional domains that are  
 CC similar to, but not identical in sequence to, the original target  
 CC functional domain. The new method enables proteins having a common  
 CC function to be identified. Identification of novel SH3 proteins will be  
 CC useful for a better understanding of cell growth, malignancy, signal

CC transduction processes, etc. New candidate drugs can be identified, and  
 CC their specificities (e.g. pharmacological activities) can be assessed  
 CC using the method of the invention.  
 SO Sequence 1734 BP; 418 A; 494 C; 526 G; 296 T;

Query Match 1.7%; Score 40; DB 35; Length 1734;  
 Best Local Similarity 100.0%; Pred. No. 2.03e-05;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1695 aaaaaaaaaaaaaaaaaaaaaaaaaagggcgccgc 1734  
 ||||||||||||||||||||||||||||||||||||  
 OY 2351 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGC 2390

RESULT 12  
 ID Q80228 standard; DNA; 3344 BP.  
 AC Q80228;  
 DT 31-AUG-1995 (first entry)  
 DE Rat NDF clone 22 DNA.  
 KW Alpha; beta; neu differentiation factor; NDF; human; rat; p185-neu;  
 KW tyrosine phosphorylation; differentiation; phenotype; proliferation;  
 KW wound; tumour; epithelial tissue; breast; stomach;  
 KW gastrointestinal disease; Barrett's oesophagus;  
 KW (non-)cystic kidney disease; inflammatory bowel disease; ds.  
 OS Rattus rattus.  
 FH Key  
 FT CDS Location/Qualifiers  
 FT 345..2255  
 FT /tag= a  
 FT /product= Rat NDF  
 PN MO9428133-A.  
 PD 08-DEC-1994.  
 PE 23-MAY-1994; U05769.  
 PR 21-MAY-1993; US-066384.  
 PA (AMGE-) AMGEN INC.  
 PI Hu S, Koeki RA, Liu N, Pierce GF, Sugarman BJ;  
 PI Wen D;  
 DR WPI: 95-022805/03.  
 DR P-PSDB: R68573.  
 PT New recombinant neu differentiation factors and corresp. DNA -  
 PT are used in the treatment of tumours, dermal wounds, and  
 PT gastrointestinal kidney and inflammatory bowel diseases.  
 PS The sequences given in Q80223-33 encode rat neu differentiation  
 CC factors (NDF's). The peptides encoded by these cDNA clones isolated  
 CC from rat tissues and cell lines, possess the ability to stimulate  
 CC human p185-neu tyrosine phosphorylation. These peptides have the  
 CC ability to induce a differentiated phenotype in certain cell lines and  
 CC can stimulate or inhibit proliferation of certain cell lines. The  
 CC NDF's can be used to treat wounds, tumours derived from epithelial  
 CC tissue of the breast, stomach etc., gastrointestinal disease, Barrett's  
 CC oesophagus, (non-)cystic kidney disease or inflammatory bowel disease.  
 CC These cDNA sequences were used in the isolation of related sequences  
 CC from human cDNA libraries which encode human NDF's.  
 SO Sequence 3344 BP; 962 A; 871 C; 749 G; 762 T;

Query Match 1.7%; Score 40; DB 13; Length 3344;  
 Best Local Similarity 100.0%; Pred. No. 2.03e-05;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3305 aaaaaaaaaaaaaaaaaaaaaaaaaagggcgccgc 3344  
 ||||||||||||||||||||||||||||||||||||  
 OY 2351 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGC 2390

RESULT 13  
 ID V02308 standard; cDNA; 3451 BP.  
 AC V02308;  
 DT 06-MAY-1998 (first entry)  
 DE Cell membrane proton-ATPase encoding cDNA.  
 KW Cell membrane proton-ATPase; Cyanidium caldarium; acid resistance;  
 KW transgenic plant; ds.  
 OS Cyanidium caldarium.  
 FH Key  
 FT Location/Qualifiers



\*\*\*\*\*  
 WIRE RELEASE  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mparc\_hm n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 19:34:43 1999; Mparc time 3738.48 Seconds  
 1512.673 Million cell updates/sec

Tabular output not generated.

Title: >US-08-946-869-9  
 Description: (1-2393) from US08946869.seq  
 Perfect Score: 2393  
 N.A. Sequence: 1 AAGCCACCTCCGCGCTCTC.....AAAAAAGCGCGCGCGGT 2393  
 Comp: TTCGGGTGAGCGCGCGAGAG.....TTTTCCTCCGCGCGCGCA

Scoring table: TABLE default  
 Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: emb157  
 1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
 7:em\_com 8:em\_or 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pi  
 13:em\_ro 14:em\_un 15:em\_vl  
 Database: genbank110  
 16:gb\_pat 17:gb\_pat2 18:gb\_htg 19:gb\_in 20:gb\_com 21:gb\_ov  
 22:gb\_pat 23:gb\_ph 24:gb\_pi 25:gb\_pi2 26:gb\_pi1  
 27:gb\_pat2 28:gb\_pat3 29:gb\_ro 30:gb\_st 31:gb\_sts 32:gb\_sy  
 33:gb\_un 34:gb\_vl

Statistics: Mean 12.262; Variance 12.056; scale 1.017

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	258	10.8	271	31	G22461 human STS WI-30695.	6.03e-89
2	118	4.9	6220	28	AB001535 Homo sapiens mRNA, com	5.59e-31
3	86	3.6	78064	18	HSAC000376 *** SEQUENCING IN PROG	2.11e-18
4	86	3.6	155074	28	AC003693 Human Chromosome 11p15	2.11e-18
5	78	3.3	7218	22	Sequence 14 from paten	2.39e-15
6	54	2.3	7218	22	Sequence 14 from paten	1.51e-06
7	53	2.3	78064	18	HSAC000376 *** SEQUENCING IN PROG	1.86e-03
8	45	1.9	148869	18	AC006037 *** SEQUENCING IN PROG	1.86e-03
9	45	1.9	193381	18	AC005538 *** SEQUENCING IN PROG	1.86e-03
10	40	1.7	1016	25	AF005158 Arabidopsis thaliana M	8.08e-02
11	40	1.7	2285	29	AF024620 Mus musculus gamma-am	8.08e-02
12	40	1.7	2325	29	RN30290 Rattus norvegicus gata	8.08e-02
13	40	1.7	2865	27	HS18167 Homo sapiens (subclone	8.08e-02

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
14	40	1.7	3451	22	E13998	8.08e-02
15	40	1.7	3519	27	HS18169	8.08e-02
16	40	1.7	6922	29	AB017793 Mus musculus mRNA for	8.08e-02
17	40	1.7	74371	27	AC005369 Homo sapiens chromosome	3.85e-02
18	40	1.7	75270	28	AF023268 Homo sapiens c1k2 kina	3.85e-02
19	41	1.7	151239	18	HS37716 Human DNA sequence ***	3.50e-01
20	38	1.6	965	22	AR024229 Sequence 22 from paten	1.69e-01
21	39	1.6	965	22	AR024229 Sequence 22 from paten	1.69e-01
22	38	1.6	1024	20	BF079414 Bos taurus common sal	3.50e-01
23	38	1.6	1285	24	ADP619 A. delticosa p91p mRNA	3.50e-01
24	38	1.6	1425	22	ADP619 Sequence 1 from patent	3.50e-01
25	38	1.6	1425	22	ADP619 Sequence 1 from patent	3.50e-01
26	38	1.6	1425	22	ADP619 Sequence 1 from patent	3.50e-01
27	38	1.6	1451	27	AF013988 Homo sapiens serine pr	1.59e-01
28	39	1.6	1659	20	SU049435 Sus scrofa growth horm	3.50e-01
29	38	1.6	1799	24	AF022082 Arabidopsis thaliana 8	3.50e-01
30	38	1.6	2399	22	ADP619 Synthetic gene for pla	3.50e-01
31	38	1.6	2573	19	DM049724 Drosophila melanogaste	3.50e-01
32	38	1.6	2775	21	PAF0AT Pseudopneumocystis ame	3.50e-01
33	38	1.6	2862	22	126126 angiotensin II type 2	3.50e-01
34	38	1.6	2862	22	126126 Human DNA sequence fro	1.69e-01
35	38	1.6	3433	27	HSN7467 Human DNA sequence fro	3.50e-01
36	39	1.6	35730	27	HSN7467 Human DNA sequence ***	1.69e-01
37	38	1.6	35730	27	HSN7467 Human DNA sequence ***	1.69e-01
38	39	1.6	151840	18	HS919B11 Human DNA sequence ***	1.69e-01
39	39	1.6	169144	28	AC005849 Homo sapiens chromosome	1.69e-01
40	40	1.6	183827	18	AC005899 Human DNA sequence ***	1.69e-01
41	40	1.6	212175	18	HS75M14 Human DNA sequence ***	1.69e-01
42	40	1.6	212175	18	HS75M14 Human DNA sequence ***	1.69e-01
43	43	1.5	130027	27	AC004982 Homo sapiens PAC clone	7.20e-01
44	44	1.5	189666	18	AC004052 *** SEQUENCING IN PROG	7.20e-01
45	37	1.5	220000	18	AC004480	7.20e-01

# ALIGNMENTS

RESULT 1 G22461 271 bp DNA STS 31-MAY-1996  
 LOCUS human STS WI-30695.  
 DEFINITION G22461  
 ACCESSION G1342787  
 NID  
 KEYWORDS STS sequence; primer; sequence tagged site.  
 SOURCE human STS derived from sequences in dbEST and the UniGene  
 collection.

ORGANISM Homo sapiens  
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrates; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE Hudson, F. Whitehead Institute/MIT Center for Genome Research; Physically  
 Mapped STS  
 TITLE Whitehead Institute/MIT Center for Genome Research; Physically  
 Mapped STS  
 JOURNAL Unpublished (1995)

COMMENT Contact: Thomas Hudson  
 Whitehead Institute/MIT Center for Genome Research  
 Whitehead Institute for Biomedical Research  
 9 Cambridge Center, Cambridge MA 02142 USA  
 Tel: 617 252 1900  
 Fax: 617 252 1902  
 Email: thudson@genome.wi.mit.edu

Primer A: CCCAGTGTGAGAAATCT  
 Primer B: CTGGAGATCAAGGCT  
 STS size: 112  
 PCR Profile:  
 Denaturation: 95 degrees C  
 Annealing: 56 degrees C  
 Polymerization: 72 degrees C  
 PCR Cycles: 35  
 Thermal Cycler:  
 Protocol: 10 ng  
 Template: 10 ng







# RELEASE

(TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MParch\_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 21:30:34 1999; Maspar time 353.20 Seconds  
Tabular output not generated. 921.792 Million cell updates/sec

Title: >US-08-946-869-9  
Description: (1-2393) from US08946869.seq  
Perfect score: 2393  
N.A. Sequence: 1 AAGGCCACCTCGCCGCTGTC.....AAAAAAGGCGCGCGCGT 2393  
Comp: TCCCGGTGAGCGCGGAGAG.....TTTTTCCCGCGCGCGCGCA

Scoring table: TABLE default

Mismatch STD: Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database:

n-geneseq32  
1:part1 2:part2 3:part3 4:part4 5:parts 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics: Mean 9.947; Variance 7.935; scale 1.253

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	116	4.8	116	22	T52991	Human gene signature
2	48	2.0	91	9	051746	Oligonucleotide probe
3	45	1.9	204	1	N81164	Base substituted E.co
4	43	1.8	91	9	051746	Oligonucleotide probe
5	42	1.8	204	1	N81164	Base substituted E.co
6	40	1.7	114	12	070468	Generic DNA sequence
7	40	1.7	172	33	T75363	Human interferin 8 a
8	40	1.7	946	13	080219	Human NDF- $\alpha$ phaz2 clone
9	40	1.7	1098	13	080225	Rat NDF clone 4 DNA
10	40	1.7	1651	13	080218	Human NDF- $\alpha$ phaz2 clone
11	40	1.7	1734	35	T39792	Human SHP9 gene
12	40	1.7	3344	13	080228	Rat NDF clone 22 DNA
13	40	1.7	3451	38	V02308	Cell membrane protein

ID	Score	Query Match	Length	DB ID	Description	Pred. No.
14	38	1.6	114	12	070466	Generic DNA sequence
15	38	1.6	178	32	T76405	Human endothelin-1 an
16	39	1.6	359	38	V02169	Human secreted protei
17	38	1.6	359	39	T88086	3' portion of cDNA c1
18	38	1.6	1435	23	T75320	Cytoplasmic anti-prote
19	39	1.6	2335	13	080216	Human PROMPF- $\alpha$ phaz2
20	39	1.6	2400	1	005055	Placenta-specific pro
21	38	1.6	2862	14	084612	Mouse A72 receptor CD
22	33	1.5	56	9	052732	Sequence of oligo nuc
23	36	1.5	114	12	070467	Generic DNA sequence
24	36	1.5	114	12	070469	Generic DNA sequence
25	36	1.5	114	12	070470	Generic DNA sequence
26	36	1.5	114	12	070467	Generic DNA sequence
27	36	1.5	162	32	T75307	Human RANTES antisens
28	36	1.5	201	38	V00418	3' fragment of clone
29	36	1.5	607	38	V02138	Human secreted protei
30	36	1.5	607	39	T97397	Human secreted protei
31	36	1.5	607	39	T88059	Partial cDNA clone en
32	36	1.5	701	32	T77001	Mouse PLAZ8 cDNA (86a
33	36	1.5	810	32	T48835	Mouse PLAZ8 cDNA (v11
34	37	1.5	857	22	T18364	Human milk kappa-case
35	37	1.5	857	8	Q46850	Human NDF-beta3 clone
36	37	1.5	906	13	080222	Rabbit zona pellucida
37	37	1.5	1338	11	055607	T. nileum GAD6H
38	37	1.5	1348	12	078278	Human MAP kinase p38-
39	36	1.5	1502	32	V10099	Chemokine superfamily
40	36	1.5	1737	17	Q99007	DNAX accessory molecu
41	37	1.5	2611	31	T67231	Rat A72 receptor clone
42	36	1.5	2884	14	084613	Sequence of murine OS
43	37	1.5	3581	10	044391	Stem cell leukaemia (
44	36	1.5	4091	1	004035	Stem cell leukaemia (
45	36	1.5	4199	1	005330	Stem cell leukaemia (

## ALIGNMENTS

RESULT 1  
ID T52991 standard; cDNA to mRNA; 116 BP.  
AC T52991;  
DT 28-OCT-1996 (first entry)  
DE Human gene signature HMG508226.  
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KW human; cloning; mapping; non-biased library; diagnosis; detection;  
OS cell typing; abnormal cell function; ss.  
OS Homo sapiens.  
PN NC0514772-A1.  
PD 01-JUN-1995.  
PE 11-NOV-1994; J01916.  
PR 12-NOV-1993; JP-355504.  
PA (MATS//) MATSUBARA K.  
PI (OKUB//) OKUBO K.  
PA Matsubara K. Okubo K.  
DR WPI; 95-206931/27.  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissues  
PS Claim 1, Page 1976; 2245p; Japanese.  
CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC given in T52991-726837 and which is able to hybridise to part of  
CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.

03-APR-1987: US-034819.  
(SUSO) SUOMEN SOKERI OY.  
Lehtovaara P, Knowles J, Kolvula A, Bamford J, Reinikkaenen T

RESULT 5  
ID N81164 standard; DNA; 204 BP.

Db 61 GTGTCCTCTTACCCAGAGACCTGACAGCTCCAGATGATACGCCCGGATCCAG 120  
 Cp 2294 GTGTCCTCTTACCCAGAGACCTGACAGCTCCAGATGATACGCCCGGATCCAG 2235  
 Db 121 GCTTATCTCTCCAGAGCTGAGTCTTCTGAGAGAGCCGGGCACTGTGTGTGA 180  
 Cp 2234 GCTTATCTCTCCAGAGCTGAGTCTTCTGAGAGAGCCGGGCACTGTGTGTGA 2175  
 Db 181 AGATGACACTCCCAAGAGTGTCTTCTGACAGTGGCCCAATGACATGGGGCCACTCA 240  
 Cp 2174 AGATGACACTCCCAAGAGTGTCTTCTGACAGTGGCCCAATGACATGGGGCCACTCA 2115  
 Db 241 AGACAAAGCCCAAGAGTGTCTTCTGACAGTGGCCCAATGACATGGGGCCACTCA 300  
 Cp 2114 AGACAAAGCCCAAGAGTGTCTTCTGACAGTGGCCCAATGACATGGGGCCACTCA 2055  
 Db 301 TCCCTGTGGGGGCTTCTTCTGAAAGTCCGACAGAGGCTCACTTTGGACCCAGCA 360  
 Cp 2054 TCCCTGTGGGGGCTTCTTCTGAAAGTCCGACAGAGGCTCACTTTGGACCCAGCA 1995  
 Db 361 GATCAGGGGTCGCGCCACCTGGGGGACAGCAGCA 400  
 Cp 1994 GATCAGGGGTCGCGCCACCTGGGGGACAGCAGCA 1955

RESULT 5  
 LOCUS AA878567 381 bp mRNA EST 25-MAR-1998  
 DEFINITION o17b06.a1 NCI-CGAP\_K1d5 Homo sapiens cDNA clone IMAGE:1492403 3'  
 mRNA sequence.  
 ACCESSION AA878567  
 NID 92987532  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 381)  
 REFERENCE 1 (bases 1 to 381)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/cgi/gap  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT  
 CONTACT: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLNC at:  
 www-bio.lnlnl.gov/db/ftp/image/image.html

FEATURES  
 source  
 Seq primer: -40m3 fwd. EST from Amersham  
 High quality sequence scop: 381.  
 Location/Qualifiers  
 1..381  
 /organism="Homo sapiens"  
 /note="Organ: kidney; Vector: p773D-PAC (Pharmacia) with  
 a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer (5'  
 AACTGGAAGATTCGCGCCGCAATATTTTCTTTTCTTTTCTTTT 3')/  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified p773 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 /db\_xref="taxon:9606"  
 /clone\_1db="NCI-CGAP\_K1d5"  
 /tissue\_type="2 pooled tumors (clear cell type)"  
 /lab\_host="DH10b"

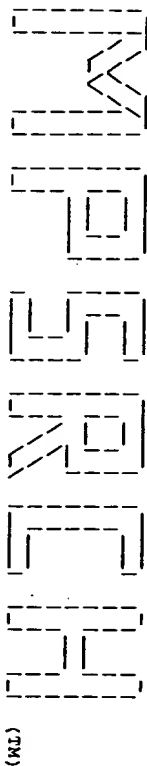
BASE COUNT 70 a 101 c 126 g 84 t  
 ORIGIN  
 Query Match 15.8%; Score 377; DB 13; Length 381;  
 Best Local Similarity 99.5%; Pred. No. 0.00e+00; Mismatches 2; Indels 0; Gaps 0;  
 Matches 379; Conservative

Db 1 TTTTTCCTCTGAAAAGGCTTATTTCCCAAGTGTGAGATCTGTGAGTGTGAGG 60  
 Cp 2358 TTTTTCCTCTGAAAAGGCTTATTTCCCAAGTGTGAGATCTGTGAGTGTGAGG 2299  
 Db 61 GTCGTGAGCTCTTACCCCAAGAGACCTCCAGCTCCAGATGATACGCCCGGAT 120  
 Cp 2298 GTCGTGAGCTCTTACCCCAAGAGACCTCCAGCTCCAGATGATACGCCCGGAT 2239  
 Db 121 CCAGGCTTGTATCTCTCCAGAGCTGTGAGTGTCTTCTGAGAGAGCCGGGCACTGTGATT 180  
 Cp 2238 CCAGGCTTGTATCTCTCCAGAGCTGTGAGTGTCTTCTGAGAGAGCCGGGCACTGTGATT 2179  
 Db 181 TGTAAAGATGACACTCCCAAGAGTGTCTCTGACAGTGGCCCAATGAGATGGGCTCAC 240  
 Cp 2178 TGTAAAGATGACACTCCCAAGAGTGTCTCTGACAGTGGCCCAATGAGATGGGCTCAC 2119  
 Db 241 CTCAAGAGCAAGGCCCAAGAGTGTCTCTGAGAGTGGGGCCAGGACCATGAGCTTACCTAGAGC 300  
 Cp 2118 CTCAAGAGCAAGGCCCAAGAGTGTCTCTGAGAGTGGGGCCAGGACCATGAGCTTACCTAGAGC 2059  
 Db 301 AAAATCCCTGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCACTTTGACCCA 360  
 Cp 2058 AAAATCCCTGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCACTTTGACCCA 1999  
 Db 361 GCGAGTCAAGGGGGTGGCGGC 381  
 Cp 1998 GCGAGTCAAGGGGGTGGCGGC 1978

RESULT 6  
 LOCUS N31660 413 bp mRNA EST 10-JAN-1996  
 DEFINITION yx69b04.r1 Homo sapiens cDNA clone 266959 5'.  
 ACCESSION N31660  
 NID 91152059  
 KEYWORDS EST.  
 SOURCE human clone-266959 primer-77 library-Soares melanocyte 2B9HM  
 human clone-266959 primer-77 library-Soares melanocyte 2B9HM  
 vector-p773D (Pharmacia) with a modified polylinker host-DH10B  
 (ampicillin resistant) Relte1-Not I Relte2-Eco RI Male. 1st strand  
 cDNA was primed with a Not I - oligo(dT) primer  
 (5'-GTTACCAATCTGAGTGTGAGGCGCGGCACTTTTCTTTTCTTTT-3')/  
 double-stranded cDNA was size selected, ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I and Eco  
 RI sites of a modified p773 vector (Pharmacia). Library  
 constructed by Bento Soares and M. Fatima Bonaldo. RNA from normal  
 foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
 Albino.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
 Sarcopterygii; Channata; Tetrapoda; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 413)  
 REFERENCE 1 (bases 1 to 413)  
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Merra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevisan, E., Waterston, R., Williamson, A., Woldmann, P. and  
 Wilson, R.  
 TITLE The WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT  
 CONTACT: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: [estewartson.wustli.edu](mailto:estewartson.wustli.edu) High quality sequence stops: 394  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Location/Qualifiers:



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_un n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 20:37:53 1999; Mpsrch time 3113.83 seconds

Tabular output not generated. 1376.223 Million cell updates/sec

Title: >US-08-946-869-9  
Description: (1-2393) from US08946869.seq  
Perfect Score: 2393  
N.A. Sequence: 1 AAGGCCACTGCGCGCTCTC.....AAAAAAGCGCGCGCGGT 2393  
Comp: TTCCGCTGACGCGCGAGAG.....TTTTTTCCCGCGCGCGCA

Scoring table: TABLE default  
Gap 6

Nmatch STD: Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-est56  
Database: genbank-est109  
3:gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13  
10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17  
14:gb-est18 15:gb-est19 16:gb-est20 17:gb-est21  
18:gb-est22 19:gb-est23 20:gb-est24 21:gb-est25 22:gb-est26  
23:gb-est27 24:gb-est28 25:gb-est29 26:gb-est30 27:gb-est31  
28:gb-est32 29:gb-est33

Statistics: Mean 12.350; Variance 5.384; scale 2.294

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	470	19.6	495	8	AA551759	nf99c01.s1 NCI_CGAP_Co	0.00e+00
2	414	17.3	441	10	AA633315	qh58f08.s1 NCI_CGAP_Co	0.00e+00
3	402	16.8	444	11	AA694490	ah33h08.s1 Soares test	0.00e+00
4	396	16.5	400	11	AA708532	z163d12.s1 Soares preg	0.00e+00
5	377	15.8	381	13	AA878567	oj17b06.s1 NCI_CGAP_K1	0.00e+00
6	379	15.8	413	19	N31660	yx69b04.r1 Homo sapien	0.00e+00
7	350	14.6	410	16	H18835	ym45d10.r1 Homo sapien	7.30e-292
8	301	12.6	311	17	A1092908	qa36c03.x1 Soares Nihm	7.78e-245
9	297	12.4	309	15	AA932133	om90h09.s1 NCI_CGAP_K1	5.32e-241
10	258	10.8	271	16	H18836	ym45d10.s1 Homo sapien	7.99e-204
11	257	10.7	294	6	AA026974	ZK01e10.s1 Soares preg	7.12e-203
12	255	10.7	299	7	AA454774	zx77b07.r1 Soares ovar	5.66e-201
13	249	10.4	259	10	AA592910	nm01f07.s1 NCI_CGAP_Co	2.80e-195

14	179	7.5	464	21	AA041682	mj05h02.r1 Soares mous	1.88e-129
15	171	7.1	321	21	W30639	mc14b05.r1 Soares mous	5.05e-122
16	169	7.1	437	21	AA036361	m174h05.r1 Soares mous	3.60e-120
17	157	6.6	173	5	R47363	Hf014.r1 Homo sapiens C	4.36e-109
18	155	6.5	361	18	A1173310	uc29c07.r1 Soares mous	3.02e-107
19	152	6.4	455	21	AA030393	m128a01.r1 Soares mous	1.73e-104
20	147	6.1	343	21	W07943	mb45e07.r1 Soares mous	6.66e-100
21	113	4.7	113	6	C00675	HUMG50008236, Human Ge	3.84e-69
22	69	2.9	252	12	AA754459	97SN1787 Rice Immature	3.96e-31
23	63	2.6	252	12	AA754459	97SN1787 Rice Immature	3.96e-31
24	57	2.4	247	12	AA754458	97SN1784 Rice Immature	3.19e-26
25	57	2.4	349	6	AA437965	ve33h02.r1 Ko mouse em	2.01e-21
26	47	2.0	247	12	AA754458	97SN1784 Rice Immature	1.00e-13
27	42	1.8	427	13	AA179789	zp53a03.s1 Stratogene	4.61e-10
28	41	1.7	74	9	AA774536	vm29e03.r1 Knowles Sol	2.39e-09
29	40	1.7	202	23	W40015	455 Mouse VM CDNA 11br	1.22e-08
30	41	1.7	216	23	AA241778	mw24e09.r1 Soares mous	2.39e-09
31	41	1.7	218	6	AA126871	m116b08.s1 Soares preg	2.39e-09
32	41	1.7	228	14	AA947054	vc33f02.r1 Barstead MP	1.22e-08
33	41	1.7	248	13	AA810322	vc33f02.r1 Barstead MP	1.22e-08
34	41	1.7	279	24	AA268052	vc33f02.r1 Barstead MP	1.22e-08
35	40	1.7	310	24	AA285478	vc33f02.r1 Barstead MP	1.22e-08
36	40	1.7	306	13	AA871093	vc33f02.r1 Barstead MP	1.22e-08
37	41	1.7	392	21	T18721	SC04H03-T7 membrane-fr	2.39e-09
38	41	1.7	464	18	AT042158	oy37a13.r1 Soares para	2.39e-09
39	41	1.7	492	11	AA717247	vp42f09.r1 Barstead mo	2.39e-09
40	41	1.7	493	17	A1121082	ue69e10.x1 Sugano mous	2.39e-09
41	41	1.7	506	18	A1182502	ue21f07.x1 Sugano mous	2.39e-09
42	40	1.7	552	15	AT038538	ox38f02.s1 Soares tota	1.22e-08
43	40	1.7	571	18	AT042336	oy37g01.x1 Soares para	1.22e-08
44	40	1.7	598	6	AA413319	AGEST00010 Anopheles g	1.22e-08
45	40	1.7	714	6	AA413321	AGEST00012 Anopheles g	1.22e-08

## ALIGNMENTS

RESULT LOCUS	1	AA551759	495 bp	mrna	EST	11-AUG-1997
DEFINITION		nf99c01.s1 NCI_CGAP_Co3	Homo sapiens	CDNA clone	IMAGE:978032, mrna	
ACCESSION		AA551759				
NID		92322011				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		1 (bases 1 to 495)				
AUTHORS		NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
JOURNAL		Tumor Gene Index				
COMMENT		Unpublished (1997)				

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1350  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,  
M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
www-bio.lnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. 27 from Amerisham  
High quality sequence stop: 385.

## FEATURES

1. 495  
Location/Qualifiers  
/organism="Homo sapiens"  
/notes="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand CDNA

was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization.

/db\_xref="taxon:9606"  
/clone\_image="928032"  
/clone\_lib="NCI-CGAP\_C03"  
/sex="pooled"  
/tissue\_type="colon"  
/lab\_host="DH10B"  
<1. ->495

BASE COUNT 89 a 145 c 162 g 99 t  
ORIGIN

Query Match 19.6%; Score 470; DB 8; Length 495;  
Best Local Similarity 98.8%; Pred. No. 0.00e+00;  
Matches 491; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

```

Db 1 TTCCCTGAAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGGTGTCTGT 60
CP 2352 TTCCCTGAAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGGTGTCTGT 2293
Db 61 GGTCCCTGTACCCCAAGAGACCTGACAGCTCCAGATGATTAAGCGCCGGGATCCAGC 120
CP 2292 GGTCCCTGTACCCCAAGAGACCTGACAGCTCCAGATGATTAAGCGCCGGGATCCAGC 2233
Db 121 CTGTATCTCTCCAGAGCTGGAGCTGTGTGGAGAGAGCGGGCATGCTGTGTGTAG 180
CP 2232 CTGTATCTCTCCAGAGCTGGAGCTGTGTGGAGAGAGCGGGCATGCTGTGTGTAG 2173
Db 181 GATGACATCCCAAGGTGGTCTCCAGAGTGGCCAGATGAGATGGGGCTCAGCTCAG 240
CP 2172 GATGACATCCCAAGGTGGTCTCCAGAGTGGCCAGATGAGATGGGGCTCAGCTCAG 2113
Db 241 GACAAGGCGACAGGTGGGGGCGGAGGCGGAGATGAGCTTACTAGAGCAAAATC 300
CP 2112 GACAAGGCGACAGGTGGGGGCGGAGGCGGAGATGAGCTTACTAGAGCAAAATC 2053
Db 301 CCCTGTGGTGGCTTCTCTTGAAGTCCGCAAGAGGCTCAGTCTTTGAGCCAGGAGG 360
CP 2052 CCCTGTGGGCTTCTCTTGAAGTCCGCAAGAGGCTCAGTCTTTGAGCCAGGAGG 1993
Db 361 TCAGGGGTGGGGGCGCCAC-TGGGG-CAGCAAGGAGAGCGGCTCAGGGGCTCGGC-ACC 417
CP 1992 TCAGGGGTGGGGGCGCCAC-TGGGG-CAGCAAGGAGAGCGGCTCAGGGGCTCGGCACC 1933
Db 418 CACCCAGAGAGCGGCTACACTGCTGAGACTCCGCTCCAGACATTTCAAGGCGCTGTTCG 477
CP 1932 CACCCAGAGAGCGGCTACACTGCTGAGACTCCGCTCCAGACATTTCAAGGCGCTGTTCG 1873
Db 478 TACTCGCGGATGTGTCC 494
CP 1872 TACTCGCGGATGTGTCC 1856

```

RESULT 2  
LOCUS AA633315 441 bp mRNA  
DEFINITION ng58f08.s1 NCI-CGAP\_C09 Homo sapiens cDNA clone IMAGE:1148103, mRNA  
ACCESSION AA633315  
NID 92556729  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 441)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck,  
M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
www-bio.lnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 433.  
Location/Qualifiers

## FEATURES

source 1.441  
/organism="Homo sapiens"  
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; 1st strand cDNA was prepared from  
RER+ colon tumor, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pT73  
vector. Library is not normalized. Library was  
constructed by Bento Soares and M. Fatima Bonaldo  
(Soares4)."  
/db\_xref="taxon:9606"  
/clone\_image="1148103"  
/clone\_lib="NCI-CGAP\_C09"  
/tissue\_type="Colon"  
/lab\_host="DH10B"  
BASE COUNT 84 a 126 c 144 g 87 t  
ORIGIN

Query Match 17.3%; Score 414; DB 10; Length 441;  
Best Local Similarity 98.0%; Pred. No. 0.00e+00;  
Matches 433; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

```

Db 1 TTTTCTGGAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGGTGTCTGT 60
CP 2355 TTTTCTGGAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGGTGTCTGT 2296
Db 61 TGTGTCCCTGTACCCCAAGAGACCTGACAGCTCCAGATGATTAAGCGCCGGGATCCA 120
CP 2295 TGTGTCCCTGTACCCCAAGAGACCTGACAGCTCCAGATGATTAAGCGCCGGGATCCA 2236
Db 121 GGCCTGTATCTCCAGAGCTGGAGCTGTCTAGAGAGAGCGGGCATGCTGTGTGT 180
CP 2235 GGCCTGTATCTCCAGAGCTGGAGCTGTCTAGAGAGAGCGGGCATGCTGTGTGT 2176
Db 181 AAGGATGACATCCCAAGAGTGGTCTGACAGTGGCCAGATGAGATGGGCTCAGCTC 240
CP 2175 AAGGATGACATCCCAAGAGTGGTCTGACAGTGGCCAGATGAGATGGGCTCAGCTC 2116
Db 241 AAGGACAGGCGCACAGTGTGGGGGCGAGGCGCCAGATGAGCTTACTAGGAGCAA 300
CP 2115 AAGGACAGGCGCACAGTGTGGGGGCGAGGCGCCAGATGAGCTTACTAGGAGCAA 2056
Db 301 ATCCCTGTGGGCAATCTCTTGAAGTCCGCGACAGAGGTTCAGTCTTTGAGACCCGCG 360
CP 2055 ATCCCTGTGGGCAATCTCTTGAAGTCCGCGACAGAGGTTCAGTCTTTGAGACCCGCG 1996
Db 361 AAGTCAGAGGTTGGGGGCGCACCTGGGGGCGAGAGCGCAAGAGCGCTCAGGCGCTCGGC 420
CP 1995 AAGTCAGAGGTTGGGGGCGCACCTGGGGGCGAGAGCGCGCTCAGGCGCTCGGC 1937
Db 421 -ACCCAGCCAGAGAGCGGCTA 441
CP 1936 CACCCAGCCAGAGAGCGGCTA 1915

```

RESULT 3

LOCUS AA694490 444 bp mRNA EST 08-JAN-1998  
 DEFINITION ah3h08.s1 Soares testis NHT Homo sapiens cDNA clone 1276883 3',  
 mRNA sequence.  
 ACCESSION AA694490  
 KEYWORDS 92855428  
 EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 444)  
 NC1-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NC1-CCAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/dbrrp/image/image.html  
 Seq primer: -40m13 fwd. ET from Amersham.  
 Location/Qualifiers  
 1. 444  
 /organism="Homo sapiens"  
 /note="Vector: pT73-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech  
 Laboratories, Inc., and primed with a Not I - oligo(dT)  
 primer [5',  
 TGTACCAATCTGAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cots, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 /db\_xref="taxon:9606"  
 /clone="1276883"  
 /clone\_1lb="Soares testis NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /lab\_host="125 c 149 g 86 c"  
 BASE COUNT 84 a 125 c 149 g 86 c  
 ORIGIN  
 Query Match 16.8%; Score 402; DB 11; Length 444;  
 Best Local Similarity 96.2%; Pred. No. 0.00e+00;  
 Matches 429; Conservative 0; Mismatches 15; Indels 2; Gaps 2;  
 Db 1 CCTCGAATAGCTTATTTTCCCGAGTGTGAGAGATCTGTGAGGAGGCTGTGG 60  
 Cp 2350 CCTCGAATAGCTTATTTTCCCGAGTGTGAGAGATCTGTGAGGAGGCTGTGG 2291  
 Db 61 TCCCTGTACCCCAAGAGACCTGACAGCTCCAGATGATTAACGGCCGAGATCCAGGCT 120  
 Cp 2290 TCCCTGTACCCCAAGAGACCTGACAGCTCCAGATGATTAACGGCCGAGATCCAGGCT 2231  
 Db 121 TGATCCTCCAGAGCTGGAGTGTCTGAGAGAGACCGGCGATCTGTGTTTGAAGA 180  
 Cp 2230 TGATCCTCCAGAGCTGGAGTGTCTGAGAGAGACCGGCGATCTGTGTTTGAAGA 2171  
 Db 181 TGACATCCCAAAAGTGTCTGTGACATGAGCCCAATGAGATGGGCTCACCTCAAGA 240  
 Cp 2170 TGACATCCCAAAAGTGTCTGTGACATGAGCCCAATGAGATGGGCTCACCTCAAGA 2111  
 Db 241 CAAGGCCACAGAGTGGGGGCGGATGGAGATGAGCTTACTAGAGCAAAATGCC 300  
 Cp 2110 CAAGGCCACAGAGTGGGGGCGGATGGAGATGAGCTTACTAGAGCAAAATGCC 2051

Db 301 CTGTGGCTGTCTCTCTTGTGAGTCCCGCAGAGGCTCAGTCTTGGACCCAGGAGGCTC 360  
 Cp 2050 CTGTGGGCTGTCTCTCTTGTGAGTCCCGCAGAGGCTCAGTCTTGGACCCAGGAGGCTC 1991  
 Db 361 AGGGGTGGCGGCC-ACCTGGGCGACATATGACAGAGGCGCTCAGAGGCTGGC-ACCCA 418  
 Cp 1990 AGGGGTGGCGGCCACCTGGGCGACAGACAGAGGCGCTCAGAGGCTGGCAGCCCA 1931  
 Db 419 CCCAGAGAGCGGCTACACTGCTGA 444  
 Cp 1930 CCCAGAGAGCGGCTACACTGCTGA 1905  
 RESULT 4  
 LOCUS AA708532 400 bp mRNA EST 24-DEC-1997  
 DEFINITION z163d12.s1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone  
 506515 3', mRNA sequence.  
 ACCESSION AA708532  
 KEYWORDS 92718450  
 NID  
 SOURCE EST.  
 ORGANISM human.  
 Homo sapiens  
 Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 400)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Joat, S.,  
 Krizman, D., Kucaba, T., Lacey, M., Le, N., Lennon, G., Marra, M.,  
 Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,  
 Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
 WashU-NCI human EST Project  
 Unpublished (1997)  
 TITLE  
 JOURNAL  
 COMMENT  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: eastewatson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 389.  
 Location/Qualifiers  
 1. 400  
 /organism="Homo sapiens"  
 /note="Organ: uterus; Vector: pT73-Pac; Site\_1: Not I;  
 Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
 oligo(dT) primer [5',  
 AACTGAGATTCGCGCGCGCTTTTCTTTTCTTTTCTTTT 3'].  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization. Library  
 constructed by M. Fatima Bonaldo."  
 /db\_xref="taxon:9606"  
 /db\_xref="GDB:381264"  
 /clone="506515"  
 /clone\_1lb="Soares pregnant uterus NBHPU"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /lab\_host="107 c 136 g 81 c"  
 BASE COUNT 76 a 107 c 136 g 81 c  
 ORIGIN  
 Query Match 16.5%; Score 396; DB 11; Length 400;  
 Best Local Similarity 99.5%; Pred. No. 0.00e+00;  
 Matches 398; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Db 1 TTTTCTCTGAAATGGCTTATTTTCCCGAGTGTGAGAGATCTGTGAGGAGGCT 60  
 Cp 2354 TTTTCTCTGAAATGGCTTATTTTCCCGAGTGTGAGAGATCTGTGAGGAGGCT 2295



Db 61 GTGTCCTCCTGTTACCCCAAGACCCCTGACCTCCAGATGATAGAGCCCGGATCCAG 120  
 Cp 2294 GTGGTCCCTGTATACCCCAAGACCCCTGACCTCCAGATGATAGAGCCCGGATCCAG 2235  
 Db 121 GCTTGATCTCTCCAGAGCTGGAGCTGTTCTGGAGAGAGCCGGCATCTGTGTTGTA 180  
 Cp 2234 GCTTGATCTCTCCAGAGCTGGAGCTGTTCTGGAGAGAGCCGGCATCTGTGTTGTA 2175  
 Db 181 AGATGACATCTCCCAAGAGTGTCTGTACATGAGCCCAAGATGATAGAGCTGATCTCA 240  
 Cp 2174 AGATGACATCTCCCAAGAGTGTCTGTACATGAGCCCAAGATGATAGAGCTGATCTCA 2115  
 Db 241 AGACAAAGCCCAAGAGTGTCTGTACATGAGCCCAAGATGATAGAGCTGATCTCA 300  
 Cp 2114 AGACAAAGCCCAAGAGTGTCTGTACATGAGCCCAAGATGATAGAGCTGATCTCA 2055  
 Db 301 TCCCTGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCAGTCTTGGAGCCAGCA 360  
 Cp 2054 TCCCTGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCAGTCTTGGAGCCAGCA 1955  
 Db 361 GGTCAAGGGGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCAGTCTTGGAGCCAGCA 400  
 Cp 1954 GGTCAAGGGGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCAGTCTTGGAGCCAGCA 1955

RESULT 5  
 LOCUS AA878567 381 bp mRNA EST 25-MAR-1998  
 DEFINITION O117D06.s1 NCI\_CGAP\_K1d5 Homo sapiens cDNA clone IMAGE:1492403 3',  
 mRNA sequence.  
 ACCESSION AA878567  
 NID 92987532  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 381)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amerisham  
 High quality sequence stop: 381.  
 Location/Qualifiers  
 1. 381

FEATURES  
 source  
 /organism="Homo sapiens"  
 /note="Organ: Kidney; Vector: p773D-Pac (Pharmacia) with  
 a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 AACGTGAGAAATTCGCGCGGCAATATTTTCTTTTCTTTT 3']  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified p773 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bonaldo.  
 /db\_xref="taxon:9606"  
 /clone\_image="1492403"  
 /clone\_11b="NCI\_CGAP\_K1d5"  
 /tissue\_type="7 pooled tumors (clear cell type)"  
 /lab\_host="DH10B"

BASE COUNT 70 a 101 c 126 g 84 t  
 ORIGIN  
 Query Match 15.8%; Score 377; DB 13; Length 381;  
 Best Local Similarity 99.5%; Pred. No. 0.00e+00;  
 Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Db 1 TTTTCTCTGAAATGCTTATTTCCCAAGTGTGAGATCTGTGATAGG 60  
 Cp 2358 TTTTCTCTGAAATGCTTATTTCCCAAGTGTGAGATCTGTGATAGG 2299  
 Db 61 GTCTGTGTCCTGTTACCCCAAGACCTGACCTCCAGATGATAGAGCCCGGAT 120  
 Cp 2298 GTCTGTGTCCTGTTACCCCAAGACCTGACCTCCAGATGATAGAGCCCGGAT 2239  
 Db 121 CCAGGCTTGATCTCCCAAGGCTGGAGCTGTTCTGGAGAGAGCCCGGATCTGTG 180  
 Cp 2238 CCAGGCTTGATCTCCCAAGGCTGGAGCTGTTCTGGAGAGAGCCCGGATCTGTG 2179  
 Db 181 TGTAGATGACACTCCCAAGGCTGCTGACAGTGGCCAGATGATAGAGGCTCAG 240  
 Cp 2178 TGTAGATGACACTCCCAAGGCTGCTGACAGTGGCCAGATGATAGAGGCTCAG 2119  
 Db 241 CTCAGAGACAGGCCACAGAGTGGGGGCGGAGGCCAGATGATAGAGCTTCTAGAGC 300  
 Cp 2118 CTCAGAGACAGGCCACAGAGTGGGGGCGGAGGCCAGATGATAGAGCTTCTAGAGC 2059  
 Db 301 AAAATCCCTGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCAGTCTTGGAGCCA 360  
 Cp 2058 AAAATCCCTGTGGGGGCTTCTCTTGAAGTCCGACAGAGGCTCAGTCTTGGAGCCA 1999  
 Db 361 GGCAGGTCAAGGGGTGGCGGC 381  
 Cp 1998 GGCAGGTCAAGGGGTGGCGGC 1978

RESULT 6  
 LOCUS N31660 413 bp mRNA EST 10-JAN-1996  
 DEFINITION YX69D04.r1 Homo sapiens cDNA clone 266959 5'.  
 ACCESSION N31660  
 NID 91152059  
 KEYWORDS EST.  
 SOURCE human clone-266959 primer-T7 library-Soares melanocyte 2BHM  
 vector-p773D (Pharmacia) with a modified polylinker host-DH10B  
 (ampicillin resistant) Raltei-Not I Raltei-Eco RI Male. 1st strand  
 cDNA was primed with a Not I - oligo(dT) primer  
 [5'-TGTTACCAATCTGAAAGTGGAGCGCGGCTTTTCTTTTCTTTT-3']  
 double-stranded cDNA was size selected, ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I and Eco  
 RI sites of a modified p773 vector (Pharmacia). Library  
 constructed by Bento Soares and M. Fatima Bonaldo. RNA from normal  
 foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
 Albino.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 413)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1995)  
 COMMENT  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: eastwason.wustl.edu  
High quality sequence stops: 294

Source: IMAGE Consortium, LNLN  
This clone is available royalty-free through LNLN; contact the  
IMAGE Consortium (infoimage.lnl.gov) for further information.

## FEATURES

Location/Qualifiers

1. 413  
/organism="Homo sapiens"  
/clone="266959"

BASE COUNT 82 a 124 c 132 g 73 t 2 others

## ORIGIN

Query Match 15.8%; Score 379; DB 19; Length 413;  
Best Local Similarity 97.3%; Pred. No. 0.00e+00;  
Matches 403; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

Db 1 TAGGAGACAGCGGGAGAGCGAGCTCCGAGCCTCTGAAAGCGCAGCTCCGAGAGTGGACTT 60  
Oy 1780 TAGGAGACAGCGGGAGAGCGAGCTCCGAGCCTCTGAAAGCGCAGCTCCGAGAGTGGACTT 1839  
Db 61 GGCACTGAACAGCTGGGACATCCGAGTACGAAAGCGCGTGAAGTGTGGAGCG 120  
Oy 1840 GGCACTGAACAGCTGGGACATCCGAGTACGAAAGCGCGTGAAGTGTGGAGCG 1899  
Db 121 GGAGGTCCA-CAGTGTAGCCCGCTGGGGTGGGTGGGCGGAGGCC-TGAGCCGCTCTGC 178  
Oy 1900 GGAGGTCCAGAGTGTAGCCCGCTGGGGTGGGTGGGCGGAGGCCCGCTCTGC 1959  
Db 179 CTCTCTGCCCCCAGAGTGGGCGCCAGCCCTGACCTGCTGGTGCAGAAAGTACAGCCCT 238  
Oy 1960 CTCTCTGCCCCCAGAGTGGGCGCCAGCCCTGACCTGCTGGTGCAGAAAGTACAGCCCT 2019  
Db 239 GCTGGCGACTCTCAAGAGGAGAGCCCGCAGAGGATTTTGTCTCTAGAGTAAAGCTCATC 298  
Oy 2020 GCTGGCGACTCTCAAGAGGAGAGCCCGCAGAGGATTTTGTCTCTAGAGTAAAGCTCATC 2079  
Db 299 TGGGCTCTGGGCGAGCGAGCTGTGGCTTGTCTTGTAGTGTAGAGCCCATTCATCTGG 358  
Oy 2080 TGGGCTCTGGGCGAGCGAGCTGTGGCTTGTCTTGTAGTGTAGAGCCCATTCATCTGG 2139  
Db 359 GCCACTGTCAAGAACACTTGGAGTGTCTTGTACAAAGCAGCATGCC 412  
Oy 2140 GCCACTGTCAAGAACACTTGGAGTGTCTTGTACAAAGCAGCATGCC 2192

RESULT 7 H18935 410 bp mRNA EST 29-JUN-1995

LOCUS ym45d10.r1 Homo sapiens CDNA clone 51262 5'  
DEFINITION H18935  
ACCESSION 5885075  
NID  
KEYWORDS EST.  
SOURCE human clone-51262 library-Soares infant brain JMB vector-Lafmid BA host-DH10B (ampicillin resistant) primer-H13P1 Ralzel-Not I site2-Hind III whole brain from a 73 days post natal female. 1st strand CDNA was primed with a Not I - oligo(dT) primer [5' AACTGAGAAATTCGCGCCGAGAGATTTTCTTTTCTTTT 3']; double-stranded CDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 410)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Matra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT

GDB: G00-424-072  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 266 1800  
Fax: 314 266 1810  
Email: eastwason.wustl.edu

High quality sequence stops: 367  
Source: IMAGE Consortium, LNLN  
This clone is available royalty-free through LNLN; contact the  
IMAGE Consortium (infoimage.lnl.gov) for further information.

## FEATURES

Location/Qualifiers

1. 410  
/organism="Homo sapiens"  
/clone="51262"

BASE COUNT 59 a 130 c 127 g 90 t 4 others

## ORIGIN

Query Match 14.6%; Score 350; DB 16; Length 410;  
Best Local Similarity 94.6%; Pred. No. 7.10e-292;  
Matches 384; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

Db 1 TGGCGNAGACCTGAGGGAGGTGTGCGAGCTTCCAGCGGGGCCCCGGGCTGAGCC 60  
Oy 856 TGGCGNAGACCTGAGGGAGGTGTGCGAGCTTCCAGCGGGGCCCCGGGCTGAGCC 914  
Db 61 CATGCTCTAGAGCGAGCGGCTGC-CTNCTACTGCGCGAGCTGGAACAGAGGCA 119  
Oy 915 CAT-GCTCTAGAGCGAGCGGCTGCCTTCTCTGCGAGCTGGAACAGAGGCA 973  
Db 120 CTAAGTGTCTACAGTCTCTCTCTGCGGTGGGCTGCGGCTGACCCGGGTTGTA 179  
Oy 974 CTAAGTGTCTACAGTCTCTCTCTGCGGTGGGCTGCGGCTGACCCGGGTTGTA 1033  
Db 180 CCACCTGGGCGGAGCTGTCTCTGATCGACTTCATGTTTCACGGTGGGCTGCTTA 239  
Oy 1034 CCACCTGGGCGGAGCTGTCTCTGATCGACTTCATGTTTCACGGTGGGCTGCTTA 1093  
Db 240 CATCTTCACGCTAACAAACAGCTGGGGCCCAAGATCGTATGTAGCAAGATATGAA 299  
Oy 1094 CATCTTCACGCTAACAAACAGCTGGGGCCCAAGATCGTATGTAGCAAGATATGAA 1153  
Db 300 GGAGGTGTCTCTCTCTCTCTCTCTGCGGCTGGTGGTGTAGTATGTTGGCCAC 359  
Oy 1154 GGAGGTGTCTCTCTCTCTCTCTCTCTCTGCGGCTGGTGGTGTAGTATGTTGGCCAC 1213  
Db 360 GGAGGGTCTCTGAGGCCAGCGGAGAGTACTTCCCAAGTATCTCT 405  
Oy 1214 GGAGGGTCTCTGAGGCCAGCGGAGAGTACTTCCCAAGTATCTCT 1259

RESULT 8 A1092908 311 bp mRNA EST 18-AUG-1998

LOCUS q36c03.x1 Soares\_NHMPu\_S1 Homo sapiens CDNA clone IMAGE:1688036  
DEFINITION 3', mRNA sequence.  
ACCESSION A1092908  
NID 93431884  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 311)

REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550  
 Email: Robert\_Strausberg@nih.gov  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: -40m13 fwd. RT from Amersham  
 High quality sequence stop: 272.  
 Location/Qualifiers  
 1. 311

/organism="Homo sapiens"  
 /note="Organ: mixed (see below); Vector: p773D-Pac  
 (Pharmacia) with a modified polylinker; Site 1: Not I;  
 Site 2: Eco RI. Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NbH, pregnant uterus  
 NBHPU, and fetal heart NBH19M) were mixed, and as circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-269223,  
 340488-345479, and 484488-489479."  
 /db\_xref="taxon:9606"  
 /clone\_lib="Scorae-NHMPU.S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"

BASE COUNT 65 a 81 c 93 g 72 t  
 ORIGIN

Query Match 12.4%; Score 301; DB 17; Length 311;  
 Best Local Similarity 98.4%; Pred. No. 7.78e-245;  
 Matches 306; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 TTTTCTCTCTGGAATGGCTTATTTCCCGAGTGTGAGGAATCTGTGAGTGTGAGG 60  
 Cp 2359 TTTTCTCTCTGGAATGGCTTATTTCCCGAGTGTGAGGAATCTGTGAGTGTGAGG 2360  
 Db 61 GGCTGTGTGCTCTGTATACCCCAAGAGCCCTGACGCTCCAAATGGATAGCGCCCGGA 120  
 Cp 2239 GGCTGTGTGCTCTGTATACCCCAAGAGCCCTGACGCTCCAAATGGATAGCGCCCGGA 2240  
 Db 121 TCCAGGCTGTATCTCTCCCAAGCTGTGAGTGTCTGTGAGAGCCCGGATGCTGTGT 180  
 Cp 2239 TCCAGGCTGTATCTCTCCCAAGCTGTGAGTGTCTGTGAGAGCCCGGATGCTGTGT 2180  
 Db 181 TTGAAGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCAAATGAGATGGGGCTCA 240  
 Cp 2179 TTGAAGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCAAATGAGATGGGGCTCA 2120  
 Db 241 CCTCAAGACAGGCCACAGGTGGGGGGCGAGGCCCAATGAGATGAGCTTACTTTAGAG 300  
 Cp 2119 CCTCAAGACAGGCCACAGGTGGGGGGCGAGGCCCAATGAGATGAGCTTACTTTAGAG 2060  
 Db 301 CAAAATCCCT 311  
 Cp 2059 CAAAATCCCT 2049

RESULT 9  
 LOCUS AA932133 309 bp mRNA  
 DEFINITION OM90H09.s1 NCI-CGAP\_K1d3 Homo sapiens cDNA clone IMAGE:1554497 3'  
 ACCESSION AA932133  
 NID 93086446  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 309)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT

Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert\_Strausberg@nih.gov  
 Tissue Procurement: Christopher Koskajuk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www.dio.lnl.gov/db/rt/image/image.html

Insert Length: 394 Std Error: 0.00  
 Seq primer: -40m13 fwd. RT from Amersham  
 High quality sequence stop: 309.  
 Location/Qualifiers  
 1. 309

FEATURES  
 SOURCE

/organism="Homo sapiens"  
 /note="Organ: kidney; Vector: p773D-Pac (Pharmacia) with  
 a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer,  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified p773 vector. mRNA  
 of normalization. Library constructed by Bento Soares and  
 M. Fatima Bonaldo."  
 /db\_xref="taxon:9606"  
 /clone\_lib="NCI-CGAP\_K1d3"  
 /lab\_host="DH10B"

BASE COUNT 61 a 80 c 102 g 66 t  
 ORIGIN

Query Match 12.4%; Score 297; DB 15; Length 309;  
 Best Local Similarity 98.1%; Pred. No. 5.22e-241;  
 Matches 303; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 TTTCTCTGAAATGGCTTATTTCCCGAGTGTGAGGAATCTGTGAGTGTGAGG 60  
 Cp 2353 TTTCTCTGAAATGGCTTATTTCCCGAGTGTGAGGAATCTGTGAGTGTGAGG 2294  
 Db 61 TGTCTCTGTATACCCCAAGAGCCCTGACGCTCCAAATGGATAGCGCCCGGA 120  
 Cp 2239 TGTCTCTGTATACCCCAAGAGCCCTGACGCTCCAAATGGATAGCGCCCGGA 2234  
 Db 121 CCTGATCTCCAGAGCTGGAGCTGTCTGTGAGAGAGCCCGGATGCTGTGTGA 180  
 Cp 2233 CCTGATCTCCAGAGCTGGAGCTGTCTGTGAGAGAGCCCGGATGCTGTGTGA 2174  
 Db 181 GGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCAAATGAGATGGGGCTCA 240  
 Cp 2173 GGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCAAATGAGATGGGGCTCA 2114  
 Db 241 GGACAGGCCACAGGTGGGGGGCGAGGCCCAATGAGATGAGCTTACTTTAGAG 300  
 Cp 2113 GGACAGGCCACAGGTGGGGGGCGAGGCCCAATGAGATGAGCTTACTTTAGAG 2054  
 Db 301 CCCCTGTGG 309  
 Cp 2053 CCCCTGTGG 2045

RESULT 10  
 LOCUS H18836 271 bp mRNA  
 DEFINITION YM45d10.s1 Homo sapiens cDNA clone 51262 3'  
 ACCESSION H18836  
 NID 9885076  
 KEYWORDS EST.  
 SOURCE human clone-51262 library-Scorae infant brain INIB vector-Lafield BA  
 host-DH10B (ampicillin resistant) primer-Promega -31m13 Raitel-Not

I Rajeez-Hind III Whole brain from a 73 days post natal female. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAATATGCGGCGCCGAGAAATTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lambda BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

**ORGANISM**  
Homo sapiens

**REFERENCE**  
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 271)

**AUTHORS**  
Hillier, L., Clark, N., Dubuque, T., Elliston, R., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevas, E., Waterston, R., Williamson, A., Woldmann, P. and Wilson, R.

**TITLE**  
The Washu-Merck EST Project

**JOURNAL**  
Unpublished (1995)

**COMMENT**  
GDB: 600-424-072  
Contact: Wilson RK  
Washu-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 162  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

**FEATURES**  
Location/Qualifiers  
1..271

**BASE COUNT**  
50 a 64 c 86 g 67 t 4 others

**ORIGIN**  
/organism="Homo sapiens"  
/clone="51262"

Query Match 10.8%; Score 258; DB 16; Length 271;  
Best Local Similarity 97.8%; Pred. No. 7,99e-204;  
Matches 265; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

**Db** 1 TTTTTCCTCTGTAATGGCTTTATTTCCAGTGTGAATCGTAGAGGTGA 60  
CP 2361 TTTTTCCTCTGTAATGGCTTTATTTCCAGTGTGAATCGTAGAGGTGA 2302

**Db** 61 GGGGCTGTGTCCTGTACCCCAAGAGACCTGCAGCTCCAGATGATAAGCGCCGG 120  
CP 2301 GGGGCTGTGTCCTGTACCCCAAGAGACCTGCAGCTCCAGATGATAAGCGCCGG 2242

**Db** 121 GATCAGGCTGTGATCCCTCCCAAGAGTGGTCTGAGAGAGCGGCGATGCTGTG 180  
CP 2241 GATCAGGCTGTGATCCCTCCCAAGAGTGGTCTGAGAGAGCGGCGATGCTGTG 2182

**Db** 181 GTTGTAAAGATGACACCTCCCAAGAGTGGTCTGAGAGAGCGGCGATGCTGTG 240  
CP 2181 GTTGTAAAGATGACACCTCCCAAGAGTGGTCTGAGAGAGCGGCGATGCTGTG 2122

**Db** 241 CACCTNAAGACAAGGCGGACCAAGATGGG 271  
CP 2121 CACCTNAAGACAAGGCGGACCAAGATGGG 2092

**RESULT** 11  
**LOCUS** AA026974 294 bp mRNA EST 09-MAY-1997  
**DEFINITION** 2k010.01 Soares pregnant uterus NBHPV Homo sapiens cDNA clone  
**ACCESSION** AA026974  
**NID** 91493165  
**KEYWORDS** EST.

**SOURCE**  
human.

**ORGANISM**  
Homo sapiens

**REFERENCE**  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 294)

**AUTHORS**  
Hillier, L., Clark, N., Dubuque, T., Elliston, R., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevas, E., Waterston, R., Williamson, A., Woldmann, P. and Wilson, R.

**TITLE**  
The Washu-Merck EST Project

**JOURNAL**  
Unpublished (1995)

**COMMENT**  
Contact: Wilson RK  
Washu-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 212.

**FEATURES**  
Location/Qualifiers  
1..294

**BASE COUNT**  
53 a 70 c 88 g 78 t 5 others

**ORIGIN**  
/organism="Homo sapiens"  
/note="Organ: uterus; Vector: p773-Pac; Site: 1; Not I; Site: 2; Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAATATGCGGCGCCGAGAAATTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone="469290"  
/clone="1D"-Soares pregnant uterus NBHPV"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10s"  
complement(<1..>294)  
/db\_xref="GDB:3753292"

Query Match 10.7%; Score 257; DB 6; Length 294;  
Best Local Similarity 98.0%; Pred. No. 7.12e-203;  
Matches 287; Conservative 0; Mismatches 1; Indels 5; Gaps 5;

**Db** 1 TTTTTCCTCTGTAATGGCTTTATTTCCAGTGTGAATCGTAGAGGTGA 60  
CP 2373 TTTTTCCTCTGTAATGGCTTTATTTCCAGTGTGAATCGTAGAGGTGA 2314

**Db** 61 TGTAGTGTGAGGCTGTGTCCTGTACCCCAAGAGACCTGCAGCTCCAGATG 120  
CP 2313 TGTAGTGTGAGGCTGTGTCCTGTACCCCAAGAGACCTGCAGCTCCAGATG 2254

**Db** 121 ATACGCGCGGAGATCCAGGCTTGATCTCCAGAGCTGGAGTNGTTCTGAGAGAC 180  
CP 2253 ATACGCGCGGAGATCCAGGCTTGATCTCCAGAGCTGGAGTNGTTCTGAGAGAC 2195

**Db** 181 CGGCGATGCTGTGATGTTGTAAAGATGACACCTCCCAAGAGTNGTCTGACANGTGCCA 240  
CP 2194 CGGCGATGCTGTGATGTTGTAAAGATGACACCTCCCAAGAGT-GGTCTGACA-GTGGCCCA 2137

**Db** 241 NATGACATGGGCTCAGCTCAAGACAAGAGGCGACAGGTGGGGGGCGCA 293  
CP 2136 GATGACATGGGCTCAGCTCAAGACAAGAGGCGACAGGTGGGGGGCGCA 2086

	RESULT	12	AA454774	299 bp	mRNA	EST	06-JUN-1997	
	LOCUS		Z77807.r1 Soares ovary tumor NBH07 Homo sapiens cDNA clone J09749					
	DEFINITION		5', mRNA sequence.					
	ACCESSION		AA454774					
	NID		G2177550					
	KEYWORDS		EST.					
	SOURCE		human.					
	ORGANISM		Homo sapiens					
			Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;					
			Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;					
			Homo.					
	REFERENCE		1 (bases 1 to 299)					
	AUTHORS		Allier,L., Allen,M., Bowles,L., Dubucque,T., Geisell,G., Joshi,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marie,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Trelsing,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.					
	TITLE		Mashu-Merck EST Project 1997					
	JOURNAL		Unpublished (1997)					
	COMMENT		Contact: Wilison R. Mashu-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu  This clone is available royalty-free through LLNL ; contact the IMAGE Consortium ( <a href="#">info@image.llnl.gov</a> ) for further information. Putative full length read The vector to vector length is 504 Seq primer: -28mj rev2 ET from Amershams High quality sequence stop: 288.  Location/Qualifiers 1..299 /organism="Homo sapiens" /note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA was primed with a Not I - 0.1lg0(ctd) primer [5' GTTTCACTATCGAAGTGGAAGGCGGGCCGGTTTTTTTTTTTTTTT 3'] , double-stranded CDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTZ19 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. /db_xref=taxon:9606" /cclone.lib="Soares ovary tumor NBH07" /cloned.by="809749" /tissue_type="Ovarian tumor" /lab_host="DH10B (ampicillin resistant)" <!--> /db_xref=GDB:6039468"					
B	BASE COUNT		51 A     95 C                 99 G                 54 T					
O	ORIGIN							
D			mRNA					
C			<--> Query Match                    10.7%; Score 255; DB 7; Length 299; Best Local Similarity 95.4%; Pred. No. 5,686-201; Matches 289; Conservative 0; Mismatches 10; Indels 4; Gaps 4;					
D	Db	1	CTGAACACGTGGGCACACTCCGCCAGTAGCACAACAGCGCCCTGAAGAAGTCGAGACGGGAG 60					
Oy	OY	1844	CTAAAACACTGGGACACATTCCCGCAGTAGCACAGCCCCTGAAGAAGTCGAGACGGGAG 1903					
D	Db	61	GTCAGC-ACTGTAGCGCGCTCTGTGGGTGGGTGGGCGGACCCTGAGCCGATCGGCTTG 119					
Oy	OY	1904	GTCACACAGTAGTAGCGCGCTCTGTGGGTGGGTGGGCGGACCCTGAGCCGCTTGCCTTG 1963					
D	Db	120	CTCCCCCAAGTGGGCGCGGCACCC-C-TAACCTGCCTGGGTCCAAGAGCATGAGCCCTGCTG 178					
Oy	OY	1964	CTCCCCCAAGTGGGCGCGGCACCCCTTAACCTGCCTGGGTCCAAGAGCATGAGCCCTGCTG 2023					
F		179	GC GGACTTC AAGAGAGAGAGAGCC - AC AGG - AT TT TG CT CCT CA GA GTA AG GC CAT AT CG G 236					

QY	2024	GGGGAATTCMAAGAAAGCCCGACAGGGGATTTGGCTTAGAGTAAGGCTCATCTGGG	2083
Db	237	CCCTGGCCCCCGACCGATGGTGGCTTGTCTCTGAGGTGAGCCCATGTGCATCTGGGCCA	296
QY	2084	CCTCGGCCCCCGACACTGGTGGCTTGTCTCTGAGGTGAGCCCATGTGCATCTGGGCCA	2143
Db	297	CTG 299	
QY	2144	CTG 2146	
RESULT	13		
LOCUS	AA592910	259 bp	mRNA
DEFINITION	nm10707.81 NCI_CGAP_C09 Homo sapiens cDNA clone IMAGE:1076485,	EST	25-SEP-1997
ACCESSION	AA592910		
NID	92408672		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS	Primates; Carnivora; Homnidae; Homo.		
TITLE	1 (bases 1 to 259)		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Tel: (301) 496-1550		
	Email: Robert_Strausberg@nih.gov		
	Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck,		
	M.D., Ph.D.		
	cDNA Library Preparation: M. Bento Soares, Ph.D.		
	cDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution Information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	www.bio.llnl.gov/dbip/image/image.html		
	Insert Length: 233 Std Error: 0.00		
	Seq primer: -40ml3 fwd. ET from Amersham		
	High quality sequence stop: 240.		
FEATURES	Location/Qualifiers		
source	1..259		
	/organism="Homo sapiens"		
	/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a		
	modified polylinker; 1st strand cDNA was prepared from		
	RER+ colon tumor, and was then primed with a Not I -		
	oligo(dT) primer. Double-stranded cDNA was ligated to Eco		
	Ri adaptors (Pharmacia), digested with Not I and cloned		
	into the Not I and Eco RI sites of the modified pT73		
	vector. Library is not normalized. Library was		
	constructed by Bento Soares and M. Fatima Bonaldo		
	(Soares4)."		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1076485"		
	/clone_1b="NCI_CGAP_C09"		
	/issue_type="colon tumor RER+"		
	/lab_host="DH108"		
BASE COUNT	51 a 66 c 86 g 56 t		
ORIGIN			
Query Match	10.4%	Score 249, DB 10, Length 259.	
Best Local Similarity	98.8%	Pred. No. 2,80e-195;	
Matches 257, Conservative	0; Mismatches 2; Indels 1; Gaps 1;		
Db	1	CCCTGAAATGCGTTTATTCGCCAGGTAGGAATGTGTGAGCGTGAAGGCTGTGCG	60
Cp	2350	CCCTGAAATGCGTTTATTCGCCAGGTAGGAATGTGTGAGCGTGTGCG	2291
Db	61	TCCCTGTTACCCCAAGACCTCGAGCCTCAGATGATAACGGCTGGGATCCAG-CAT	119







\*\*\*\*\*  
 MPELIT (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mparch\_n n.a. - n.a. database search, using Smith-Waterman algorithm  
 Run on: Sun Feb 14 19:34:43 1999; Maspur time 3738.48 Seconds  
 Tabular output not generated. 1512.673 Million cell updates/sec

Title: >US-08-946-869-9  
 Description: (1-2393) from US08946869.seq  
 Perfect Score: 2393  
 N.A. Sequence: 1 AAGGCCACCTCGCGCTCTC.....AAAAAAGCGCGCGCGGT 2393  
 Comp: TTCGCGTGGAGCGCGGAGAG.....TTTTTTCCGCGCGCGCGCA

Scoring table: TABLE default  
 Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: emb157  
 1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
 7:em\_com 8:em\_ov 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pi  
 13:em\_ro 14:em\_un 15:em\_vl  
 genbank110

16:gb\_ba1 17:gb\_ba2 18:gb\_htg 19:gb\_in 20:gb\_com 21:gb\_ov  
 22:gb\_pat 23:gb\_pi 24:gb\_ph 25:gb\_pi 26:gb\_pi  
 27:gb\_pi2 28:gb\_pi3 29:gb\_pi 30:gb\_pi 31:gb\_pi 32:gb\_pi  
 33:gb\_un 34:gb\_vl

Statistics: Mean 12.262; Variance 12.056; scale 1.017

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
C 1	258	10.8	271	31	G22461	human STS WI-30695.	6.03e-89
C 2	118	4.9	6220	28	AB001535	Homo sapiens mRNA, com	5.59e-31
C 3	86	3.6	78064	18	HSAC000376	*** SEQUENCING IN PROG	2.11e-18
C 4	86	3.6	155074	28	AC003693	Human Chromosome 11p15	2.11e-18
C 5	78	3.3	7218	22	I66494	Sequence 14 from paten	1.51e-06
C 6	54	2.3	7218	22	I66494	Sequence 14 from paten	1.51e-06
C 7	53	2.2	78064	18	HSAC000376	*** SEQUENCING IN PROG	3.39e-06
C 8	45	1.9	148869	18	AC006037	*** SEQUENCING IN PROG	1.86e-03
C 9	45	1.9	193381	18	AC005158	*** SEQUENCING IN PROG	1.86e-03
C 10	45	1.7	1016	25	AF024650	Arabidopsis thaliana M	8.08e-02
C 11	40	1.7	2286	29	AF024650	Mus musculus gamma-aml	8.08e-02
C 12	40	1.7	2335	29	RN030290	Rattus norvegicus gala	8.08e-02
C 13	40	1.7	2865	27	HSL81867	Homo sapiens (subclone	8.08e-02

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
C 14	40	1.7	3451	22	E13998	Candidium calidarium nr	8.08e-02
C 15	40	1.7	3519	27	HSL81869	Homo sapiens (subclone	8.08e-02
C 16	40	1.7	6922	29	MM017793	Mus musculus chromosom	8.08e-02
C 17	40	1.7	74371	27	AC005369	Homo sapiens mRNA for	8.08e-02
C 18	41	1.7	75270	28	AF023268	Homo sapiens c1k2 kina	3.85e-02
C 19	41	1.7	151239	18	HS377F16	Human DNA sequence ***	3.85e-02
C 20	38	1.6	965	22	AR024229	Sequence 22 from paten	3.50e-01
C 21	39	1.6	965	22	AR024229	Sequence 22 from paten	3.50e-01
C 22	38	1.6	1024	20	BT079414	Bos taurus common all	3.50e-01
C 23	38	1.6	1285	24	ADPGIP	A. deliiosa p91p mRNA	3.50e-01
C 24	38	1.6	1425	22	AR005195	Sequence 1 from paten	3.50e-01
C 25	38	1.6	1425	22	I81228	Sequence 1 from paten	3.50e-01
C 26	38	1.6	1425	22	I81228	Sequence 1 from paten	3.50e-01
C 27	38	1.6	1451	27	AF013988	Homo sapiens beta pr	3.50e-01
C 28	39	1.6	1650	20	SF049435	Sus scrofa growth horm	1.69e-01
C 29	38	1.6	1799	24	AF024082	Arabidopsis thaliana s	3.50e-01
C 30	38	1.6	2399	22	DM049724	Synthetic gene for pla	3.50e-01
C 31	38	1.6	2573	19	AF024082	Drosophila melanogaste	3.50e-01
C 32	38	1.6	2775	21	PAFR0AT	Pseudopleuronectes ame	3.50e-01
C 33	38	1.6	2862	22	I26126	Sequence 1 from paten	3.50e-01
C 34	38	1.6	2862	22	I26126	Sequence 1 from paten	3.50e-01
C 35	38	1.6	3433	27	HS338H10	Human DNA sequence fro	1.69e-01
C 36	39	1.6	35730	26	HSNFG9	Human DNA sequence fro	1.69e-01
C 37	38	1.6	74371	27	AC005369	Human DNA sequence fro	1.69e-01
C 38	39	1.6	151840	18	HS919811	Homo sapiens chromosom	1.69e-01
C 39	39	1.6	159144	18	AC005849	Human DNA sequence ***	1.69e-01
C 40	39	1.6	163847	28	AC005899	*** SEQUENCING IN PROG	1.69e-01
C 41	39	1.6	212175	18	HS75N14	Homo sapiens chromosom	1.69e-01
C 42	39	1.6	212175	18	HS75N14	Human DNA sequence ***	1.69e-01
C 43	37	1.5	130027	27	AC004982	Homo sapiens PAC clone	7.20e-01
C 44	37	1.5	189666	18	AC000052	*** SEQUENCING IN PROG	7.20e-01
C 45	37	1.5	220000	18	AC004480	*** SEQUENCING IN PROG	7.20e-01

## ALIGNMENTS

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
C 1	258	10.8	271	31	G22461	human STS WI-30695.	6.03e-89
C 2	118	4.9	6220	28	AB001535	Homo sapiens mRNA, com	5.59e-31
C 3	86	3.6	78064	18	HSAC000376	*** SEQUENCING IN PROG	2.11e-18
C 4	86	3.6	155074	28	AC003693	Human Chromosome 11p15	2.11e-18
C 5	78	3.3	7218	22	I66494	Sequence 14 from paten	1.51e-06
C 6	54	2.3	7218	22	I66494	Sequence 14 from paten	1.51e-06
C 7	53	2.2	78064	18	HSAC000376	*** SEQUENCING IN PROG	3.39e-06
C 8	45	1.9	148869	18	AC006037	*** SEQUENCING IN PROG	1.86e-03
C 9	45	1.9	193381	18	AC005158	*** SEQUENCING IN PROG	1.86e-03
C 10	45	1.7	1016	25	AF024650	Arabidopsis thaliana M	8.08e-02
C 11	40	1.7	2286	29	AF024650	Mus musculus gamma-aml	8.08e-02
C 12	40	1.7	2335	29	RN030290	Rattus norvegicus gala	8.08e-02
C 13	40	1.7	2865	27	HSL81867	Homo sapiens (subclone	8.08e-02

LOCUS G22461 271 bp DNA STS 31-MAY-1996  
 DEFINITION G22461  
 ACCESSION G1342787  
 NID  
 KEYWORDS STS sequence; primer; sequence tagged site.  
 SOURCE human STS derived from sequences in dbEST and the Unigene  
 collection.  
 ORGANISM Homo sapiens  
 Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 271)  
 AUTHORS Hudson, T.  
 TITLE Whitehead Institute/MIT Center for Genome Research, Physically  
 Mapped STS  
 JOURNAL Unpublished (1995)  
 COMMENT

Contact: Thomas Hudson  
 Whitehead Institute/MIT Center for Genome Research  
 Whitehead Institute for Biomedical Research  
 9 Cambridge Center, Cambridge MA 02142 USA  
 Tel: 617 252 1900  
 Fax: 617 252 1902  
 Email: thudson@genome.wi.mit.edu

Primer A: CCCGAGTGTGAGAAATCTGT  
 Primer B: CTGGAGATCAAGGCGCT  
 STS size: 112  
 PCR Profile:  
 Denaturation: 94 degrees C  
 Annealing: 56 degrees C  
 Polymerization: 72 degrees C  
 PCR Cycles: 35  
 Thermal Cycler:  
 Protocol:  
 Template: 10 ng

Primer: each 5 pm  
 dNTPs: each 4 mM  
 Tag Polymerase: 0.025 units/ul  
 Total Vol: 20 ul

Buffer:  
 MgCl2: 1.5 mM  
 KCl: 50 mM  
 Tris-HCl: 10 mM  
 pH: 9.3

FEATURES  
 Derived from dbEST (genbank accession H18836).  
 Location/Qualifiers  
 1..271

Source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="349.5 CR from top of Chr19 linkage group"  
 STS  
 primer\_bind 32..143  
 primer\_bind 32..143  
 Complement(126..143)  
 BASE COUNT 50 a 64 c 86 g 67 t 4 others  
 ORIGIN

Query Match  
 Best Local Similarity 97.8%; Score 258; DB 31; Length 271;  
 Matches 285; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Db 1 TTTTTCCTCTCAATGGCTTATTTCCCACTGAGGAATCTGTGCTGA 60  
 |||||  
 Cp 2361 TTTTTCCTCTCAATGGCTTATTTCCCACTGAGGAATCTGTGCTGA 2302  
 |||||  
 Db 61 GGGCTGTGTGCTCTGCTTACCCCAAGGACCTGACCTCCAGATGATGAGGCCGG 120  
 |||||  
 Cp 2301 GGGCTGTGTGCTCTGCTTACCCCAAGGACCTGACCTCCAGATGATGAGGCCGG 2242  
 |||||  
 Db 121 GATCCAGGCTTGTATCTCTCCAGAGTGGAGTGGTCTGTGGAGAGCCGGGCTGTG 180  
 |||||  
 Cp 2241 GATCCAGGCTTGTATCTCTCCAGAGTGGAGTGGTCTGTGGAGAGCCGGGCTGTG 2182  
 |||||  
 Db 181 GTTGTAAAGTACACTCCCAAGGAGTGTCTCTGAGAGTGGCCAGATGAGATGGGGT 240  
 |||||  
 Cp 2181 GTTGTAAAGTACACTCCCAAGGAGTGTCTCTGAGAGTGGCCAGATGAGATGGGGT 2122  
 |||||  
 Db 241 CACCTAAGGACCAAGGCCACAGTCTGGG 271  
 |||||  
 Cp 2121 CACCTAAGGACCAAGGCCACAGTCTGGG 2092  
 |||||

RESULT 2  
 LOCUS AB001335 6220 bp mRNA PRI 28-NOV-1998  
 DEFINITION Homo sapiens mRNA, complete cds.  
 ACCESSION AB001335  
 NID G3928755  
 KEYWORDS KNP3  
 SOURCE Homo sapiens 20-26weeks fetus brain tissue\_11b:human fetal brain  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominoidea; Homo.  
 1 (sites)  
 Nagamine, K., Kudoh, J., Minoshima, S., Kawasaki, K., Asakawa, S.,  
 Ito, F. and Shimizu, N.  
 Molecular cloning of a novel putative Ca2+ channel protein (TRPC7)  
 highly expressed in brain  
 nucleu 5'-STRECH CDNA library (CLONTECH) and Human Brain, caudate  
 nucleus 5'-STRECH CDNA library (CLONTECH) CDNA to mRNA.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 2 (bases 1 to 6220)  
 Shimizu, N.  
 Direct Submission  
 Submitted (28-FEB-1997) to the DDBJ/EMBL/Genbank databases.  
 Nobuyoshi Shimizu, Keio University School of Medicine, Department  
 of Molecular Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo 160,

Japan (E-mail: shimizudb.med.keio.ac.jp, tel:03-3351-2370,  
 Fax:03-3351-2370)  
 Location/Qualifiers  
 1..6220

FEATURES  
 Source  
 /organism="Homo sapiens"  
 /note="7 Caucasians"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /dev\_stage="20-26weeks fetus"  
 /map="21q22.3"  
 /tissue\_11b="human fetal brain 5'-STRECH plus cDNA library  
 (CLONTECH) and Human Brain, caudate nucleus 5'-STRECH CDNA  
 library (CLONTECH)"  
 /tissue\_type="brain"  
 446..4957  
 /gene="KNP3"  
 446..4957  
 /note="KNP3"  
 /note="similar to C.elegans hypothetical protein  
 CETO1H8.1, CECO5C12.3, CEF54D1.5, similar to trp and  
 trp-like proteins"  
 /codon\_start=1  
 /db\_xref="PID:01035682"  
 /db\_xref="PID:03928756"  
 /translation="MEPSALRKAGESEDEEGEGEGLPRVTDGMYSNLRNSSLFKS  
 RLQCPFGNNDKQESLSWIPENIKKKECVFVESKISDAKRYCCQCYTHQHLPEA  
 TKPFTFOGTDPRKHOEMPTDAPGDGVPFGLSQKRYKRVSDPTSVTHLATO  
 HMGDLPRLILSYGAKNPMKRLSIFRGLKVAQTGAMITGSGTGAKOV  
 GEAVRDSLSYSYEGELITGVATKGVHNRGLIHPTGSPFPAVILDEGSGCNLTC  
 LDSNHSRFLIVDDTHQGVPEIPRLRLEFISEQTEREGVAKIPYCVVEGGP  
 GTLTIDNATNGTTPCVVEGSGRAVIAQVAVLPVSDITISLQKLSFVEGMP  
 TTESRIYETKRIODLYRRRLTLVREGDGODVATLQALLKASRSQDFGHE  
 NMDQLAVAMNRYDARSRIPEDEQWKRSDLPRTTALISNKEFVALFLENGV  
 OLKEFVMDLILYKENDPSCLFHSKLQVLDVDEPAPAPAPAPLOMHNQVLR  
 ELIDFTOPLEPRNRHNDRLLLPVPHVKNVGVSLRSKSGHVTFMDPIRD  
 LLIAIVONREELGIIWAOSDCAIAALASKLKELSKEDTDSSEMLAAEY  
 EHRAIVGETECYRDEERAKLLRVSEANGKTCLOLAEAKMKRSVSGIGIAPIT  
 KYMGQSLSYNDGLRWRTICMLAPLLTGLISPERKLDVGTAPAAARAFPAVAV  
 FHLNLSYFALCLPAAVAVADVPVSKEDCATYLLFSLVCEEMQLYRDEPGL  
 MKRALEFSPFNKLDGAILLPAGLTLCLIPATLYLPGVILSLDPLCLRLMHF  
 TISITLPLKIIIVKMKDVEFLFLAAVWVSQVAKQAILIHNERVMDLFRGAVY  
 HSITITGQIPGYIDGVNFPNHCSPGTDPPYKPCPESDATORPAPFPMILYLLC  
 LYLTIFILILINLIANFNTPFOOVDEHTQINFORHDIIEYHGRBAPPEFILLS  
 HLOLFIRVYLTAPAKRHOKLKNLENEBAALISEBYAQTARLHVITVLAAGS  
 KIEDISNKVDAMVDLIDPLKRSSEHQAALISSEBYAQTARLHVITVLAAGS  
 SEADVPILASOKAABEPDAPGGRKTEBEGDSTHVANRLILPNCVPTFPVNERV  
 PWTEFELIDPPTIARKDAAMADPGDLEPSTIOYNAVVDGLRBRSGFPGTYVO  
 AGLPLNMGRGTGRLGRGSLSCFSPNHTLYPMTVRMRNEDGALCRSIRKMLEVLYVK  
 LPLSEHVALPGSREPEMPLRLKRLIRLEHNPSPENLKCMEYKGYNDPRTD  
 NAIETVAIVSVHFQODNDVLELNLSLHNCDSGASIRHGVDRRIPLIANHKLLOK  
 AAEFGHAY"

BASE COUNT 1248 a 1938 c 1882 g 1152 t  
 ORIGIN

Query Match  
 Best Local Similarity 4.9%; Score 118; DB 28; Length 6220;  
 Matches 423; Conservative 0; Mismatches 269; Indels 6; Gaps 4;

Db 3092 TACTTACGTACTCTGGAATTAAGTGGAGTGGCCCAATCTTGTCTGTGGCAGG 3091  
 |||||  
 Oy 945 TACTTACGTACTCTGGAATTAAGTGGAGTGGCCCAATCTTGTCTGTGGCAGG 1004  
 |||||  
 Db 3092 CTGACCTGACGACATCGCGGAGAGCTGTACCCGGGGCGGATCCTCTCTGAGC 3151  
 |||||  
 Oy 1005 GTGGGCGTGGCGGCGGAGAGCTGTACCCGGGGCGGATCCTCTCTGAGC 1064  
 |||||  
 Db 3152 TTCACTCTGTCTGCTCGCGCTCATGACATTTTACATGATAGAGCTGGGGCC 3211  
 |||||  
 Oy 1065 TTCACTCTGTCTGCTCGCGCTCATGACATTTTACATGATAGAGCTGGGGCC 1124  
 |||||  
 Db 3212 AGATATATCTTGTGAAGGAGATGAGAGAGCTTCTCTCTCTCTCTGCTGCTG 3271  
 |||||  
 Oy 1125 AGATATCTCTCTGAGAGATGAGAGAGCTTCTCTCTCTCTCTCTGCTGCTG 1184  
 |||||

Db 3372 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3331  
 1185 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1244  
 Db 3332 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3391  
 1245 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1304  
 Db 3392 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3451  
 1305 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1361  
 Db 3452 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3510  
 1362 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1420  
 Db 3511 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3570  
 1421 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1480  
 Db 3571 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3630  
 1481 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1540  
 Db 3631 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3690  
 1541 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1599  
 Db 3691 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 3728  
 1600 GTGTGGGTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1637

RESULT 3  
 LOCUS HSAC000376 78064 bp DNA 24-MAR-1997  
 DEFINITION \*\* SEQUENCING IN PROGRESS \*\* Human Chromosome 11p15.5 pac from  
 AC000376  
 NID 91905927  
 KEYWORDS HTG: HTGS\_PHASE1.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 78064)  
 Evans, G.A., Bradbury, P., Brignac, S., Bumeister, R., Davis, J.,  
 Davies, C.J., Davis, C., English, C., Fondon, T., Franklin, T.L.,  
 Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J.,  
 Hinson, S., Megarity, C., Narayanaswamy, U., Newton, J., O'Brien, K.,  
 Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J.,  
 Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T.,  
 Wilson, R. and Burbee, D.  
 TITLE  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 78064)  
 AUTHORS Evans, G.A., Bradbury, P., Brignac, S., Bumeister, R., Davis, J.,  
 Davies, C.J., Davis, C., English, C., Fondon, T., Franklin, T.L.,  
 Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J.,  
 Hinson, S., Megarity, C., Narayanaswamy, U., Newton, J., O'Brien, K.,  
 Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J.,  
 Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T.,  
 Wilson, R. and Burbee, D.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-MAR-1997) Genome Science and Technology Center,  
 University of Texas Southwestern Medical Center at Dallas, 5323  
 Harry Hines Blvd, Dallas, TX 75235-8591, USA  
 COMMENT \*\*\*  
 \*\*\* WARNING: Phase 1 High Throughput Genome Sequence \*\*\*  
 \*\*\* This sequence is unfinished. It consists of 15 contigs for  
 \*\*\* which the order is not known; their order in this record is  
 \*\*\* arbitrary. In some cases, the exact lengths of the gaps

between the contigs are also unknown; these gaps are presented  
 as runs of N as a convenience only. When sequencing is complete,  
 the sequence data presented in this record will be replaced  
 by a single finished sequence with the same accession number.

1  
 2180 contig of 2179 bp in length  
 4339 contig of 2159 bp in length  
 4339 contig of 2266 bp in length  
 6635 contig of 3195 bp in length  
 9820 contig of 2431 bp in length  
 12251 contig of 2612 bp in length  
 14863 contig of 3853 bp in length  
 18716 contig of 3853 bp in length  
 22568 contig of 4893 bp in length  
 22569 contig of 6088 bp in length  
 27462 contig of 6737 bp in length  
 33549 contig of 6737 bp in length  
 40287 contig of 7652 bp in length  
 47938 contig of 7906 bp in length  
 55844 contig of 9553 bp in length  
 55397 contig of 12667 bp in length  
 65398 Location/Qualifiers  
 1.78064  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 17050 a 22863 c 21471 g 16611 t 69 others

ORIGIN

Query Match 3.6% Score 86; DB 18; Length 78064;  
 Best Local Similarity 74.7% Pred. No. 2,11e-18;  
 Matches 130; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Db 65867 TCATCTCTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 65926  
 Cp 1321 TCATCTCTGCTGCTTCTGGGGTGGCCAGCAGCCATCTCATCAACAGACGCCGG 1262  
 Db 65927 CGGAGATCCACTCCAGCGGCGGCGTCTGTCAGCAGCGCCCTGCTGTCACCG 65986  
 Cp 1261 CGGAGATCCACTCCAGCGGCGGCGTCTGTCAGCAGCGCCCTGCTGTCACCG 1202  
 Db 65987 TAGGCTACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 66040  
 Cp 1201 TAGGCTACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1148

RESULT 4  
 LOCUS AC003693 15074 bp DNA 30-SEP-1998  
 DEFINITION Human Chromosome 11p15.5 PAC clone pDJ915f1 containing KVLG1 gene,  
 complete sequence.  
 AC003693  
 NID 93687269  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Euthera;  
 Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 15074)  
 Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basit, M.,  
 Bueltner, J., Bumeister, R., Card, P., Deslipoat, F., Dunn, J.,  
 English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G.,  
 Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N.,  
 Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S.,  
 Schageman, J., Schilling, P., Schultz, R., Stimson, S., Syed, M. and Ward, T.  
 HTGS Submission  
 TITLE  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 15074)  
 AUTHORS Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basit, M.,  
 Bueltner, J., Bumeister, R., Card, P., Deslipoat, F., Dunn, J.,  
 English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G.,  
 Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N.,  
 Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S.,  
 Schageman, J., Schilling, P., Schultz, R., Stimson, S., Syed, M., Valenzuela, D.,  
 Ward, T. and Wilson, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-DEC-1997) Genome Science & Technology Center,  
 University of Texas Southwestern Medical Center, 5323 Harry Hines

REFERENCE  
 3 (bases 1 to 155074)  
 AUTHORS  
 Blvnd, Dallas, TX 75235-8591, USA  
 Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M.,  
 Buetner, J., Butler, C., Card, P., Desalilob, F., Dunn, J.,  
 English, C., Ethridge, S., Garner, H.R., Gee, Y., Gordon, M., Gotway, G.,  
 Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, R., Loo, K.N.,  
 Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S.,  
 Schageman, J., Schultz, R.A., Stimson, S., Waller, K. and Ward, T.  
 Direct Submission  
 Submitted (30-SEP-1998) Genome Science & Technology Center,  
 University of Texas Southwestern Medical Center, 5323 Harry Hines  
 Blvd, Dallas, TX 75235-8591, USA  
 On Oct 1, 1998 this sequence version replaced g1:3264564.  
 Further information regarding the map of this region or  
 annotation of PDJ915f1 can be found at  
<http://gdb.csh.mcgill.edu/chromoso.htm>  
 IMPORTANT: This submission contains the entire insert of clone  
 PDJ915f1. PDJ915f1 comes from the RPCI-3 PAC library constructed  
 at the Roswell Park Cancer Institute by the Plaster de Jong group.  
 This clone has been finished according to strict quality criteria  
 and attempts have been made to resolve all base calling problems  
 such as compressions and repetitive elements. The expected  
 phred/phrap calculated error/10kb is 0.97. In addition, attempts  
 have been made to assure over 99% of consensus base calls consist  
 of either double-stranded coverage or 2 types of labeling chemistry  
 on one strand.

CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p15.5  
 Wilms tumor type 2 (WT2) region between ESTs CTSD and IGF2(INS12).  
 This region spanning approximately 2 kbp is mapped between D1S2701  
 and D1S1145.

MARKER INFORMATION: STS WI-9248.

## FEATURES

## SOURCE

1. 155074  
 Location/Qualifiers

repeat_region	/organism="Homo sapiens"	complement(68934..69046)
repeat_region	/db_xref="taxon:9606"	/rpt_family="Alu"
repeat_region	/rpt_family="Alu"	complement(3043..3195)
repeat_region	complement(12602..12763)	/rpt_family="MER4"
repeat_region	complement(13184..13415)	/rpt_family="Alu"
repeat_region	complement(26921..27026)	/rpt_family="Alu"
repeat_region	32140..32471	/rpt_family="Alu"
repeat_region	34325..34628	/rpt_family="Alu"
repeat_region	39386..39469	/rpt_family="Alu"
repeat_region	39630..39771	/rpt_family="MLT1"
repeat_region	44752..44912	/rpt_family="MLT1"
repeat_region	45325..45410	/rpt_family="THE1"
repeat_region	complement(49939..50077)	/rpt_family="THE1"
repeat_region	52802..55183	/rpt_family="Alu"
repeat_region	53582..53663	/rpt_family="L1"
repeat_region	complement(56684..56986)	/rpt_family="Alu"
repeat_region	complement(57763..57988)	/rpt_family="Alu"
repeat_region	complement(59538..59895)	/rpt_family="L1"
repeat_region	complement(66135..66214)	/rpt_family="TAR1"
repeat_region	complement(66228..66434)	/rpt_family="MER1"
repeat_region	68089..68377	/rpt_family="MER1"
repeat_region	/rpt_family="Alu"	

repeat_region	complement(68934..69046)	/rpt_family="Alu"
repeat_region	69637..72523	/rpt_family="L1"
repeat_region	complement(70090..70376)	/rpt_family="Alu"
repeat_region	71058..71304	/rpt_family="Alu"
repeat_region	71629..71964	/rpt_family="Alu"
repeat_region	complement(80143..80239)	/rpt_family="Alu"
repeat_region	complement(83167..83415)	/rpt_family="Alu"
repeat_region	complement(84253..84559)	/rpt_family="Alu"
repeat_region	complement(84663..85133)	/rpt_family="Alu"
repeat_region	90644..90990	/rpt_family="THE1"
repeat_region	complement(91700..91966)	/rpt_family="Alu"
repeat_region	92586..92737	/rpt_family="Alu"
repeat_region	92944..93169	/rpt_family="Alu"
repeat_region	complement(93315..93666)	/rpt_family="MER42"
repeat_region	complement(93707..95255)	/rpt_family="THE1"
repeat_region	complement(95260..95632)	/rpt_family="MSR"
repeat_region	95968..96013	/rpt_family="THE1"
repeat_region	96486..97004	/rpt_family="MER42"
repeat_region	complement(98550..98704)	/rpt_family="MER1"
repeat_region	complement(99173..99442)	/rpt_family="MER42"
repeat_region	complement(99456..99916)	/rpt_family="Alu"
repeat_region	complement(10086..103880)	/rpt_family="MER42"
repeat_region	complement(101541..101608)	/rpt_family="L1"
repeat_region	complement(105454..105672)	/rpt_family="THE1"
repeat_region	complement(105673..105983)	/rpt_family="MER43"
repeat_region	complement(105997..106047)	/rpt_family="Alu"
repeat_region	complement(107379..114158)	/rpt_family="MER43"
repeat_region	111143..111419	/rpt_family="L1"
repeat_region	complement(11213..112455)	/rpt_family="Alu"
repeat_region	115302..116008	/rpt_family="MER25"
repeat_region	116348..119279	/rpt_family="L1"
repeat_region	119474..119733	/rpt_family="L1"
repeat_region	125016..125597	/rpt_family="Alu"
repeat_region	129543..139819	/rpt_family="Alu"
repeat_region	complement(130188..130341)	/rpt_family="Alu"
repeat_region	134446..134618	/rpt_family="L1"
repeat_region	complement(138186..138488)	/rpt_family="Alu"









```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

REFERENCE 1 (bases 1 to 3519)

## AUTHORS

Martin, C.H., Arcalena, T.T., Bondoc, M.M., Chang, A., Citz, P.A., Davis, C.A., Doyle, C.M., Ellison, C.L., Fardon, D.E., Gunning, K.M., Houston, K.A., Jaklevic, M.A., Kacher, K.E., Kim, K., Kim, S.F., Ko, C.L., Lewis, K.D., Li, M., Lindquist, R.J., Lomoten, M.A., Lustre, V.M., Mechuris, M.U., Mayeda, C.A., Mguen, T.M., Miller, C.A., Mok, M.S., Pacלב, J.M., Patel, S.G., Santos, R.F., Subramanian, S., Wan, K.H., Wittelev, K.R., Yee, A., Yeh, R.T., Yu, C. and Palazzolo, M.J. Sequencing of human chromosome 5q (unpublished, 1996).

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

2 (bases 1 to 3519)